Alaska Alternate Assessment

Writing Scoring Accuracy Project
Phase 4 Final Report

June 30, 2013

Table of Contents

Background

Phase 4 of the three-year *Writer Rating Study Final Plan* required that Dillard Research Associates (DRA) conduct a follow-up writing scoring accuracy study in the spring of 2013 for grade 5 writing assessments for the Alaska Alternate Assessment (AKAA). Districts who assessed 5th grade students with the AKAA in writing were identified and provided with shipment instructions and materials to return both scoring protocols and student materials to DRA for analysis. An expert rater reviewed and scored each submission from the field in the areas of Correct Word Sequences (CWS), Correct Letter Sequences (CLS), and Ideas & Organization (I & O). As CWS and I & O were studies last year, the focus of this year's study was CLS. The expert rater reviewed and scored all items, also reviewing test documents for procedural expectations and content. Anecdotal information regarding potential training topics was collected and is included in this report.

In addition, data analyses were conducted to address agreement between the Alaska Qualified Assessor and the DRA expert reviewer.

Demographics

Nineteen districts were involved in the Phase 4 writing scoring accuracy study. A total of 77 submissions were received by DRA. Twenty submissions proceeded to the Extended Levels of Support (ELOS) tasks. The accuracy of ELOS task ratings could not be evaluated, as each task related to the level of support provided during test administration. The DRA expert reviewer thus determined only if the Assessor had entered all scores for students who participated in ELOS in the scoring protocols. Only one score was missing, yet that score was entered in the online data entry system. The error thus did not impact the student.

One Assessor did not score a student's correct word sequences (CWS) task on the scoring protocol. DRA checked the AYP file to see that the student had received a score for the task, but it was not possible to determine where/how the points for the specific task should have been attributed. This task was thus not double-scored to check for rater accuracy.

Two submissions did not include a scoring protocol and thus could not be included. Therefore, a total of 75 records were included in the review. The sample analyzed thus represents 97% of possible participants in the grades reviewed (75/77).

Report Coding

The following abbreviations for item type coding listed in the remainder of the report are represented below.

Abbreviation	Description
CWS	Correct Word Sequences items
CLS	Correct Letter Sequences items
IO	Ideas and Organization ratings
r	Raw score awarded by the rater
pct	Percentage correct calculation recorded by the rater
pts	Points determined appropriate by the rater
num	The numerator in CWS items, corresponding to the number of correct word sequences the rater determined
den	The denominator in CWS items, corresponding to the maximum number of points the rater deemed possible
r1 or R1	Rater 1 - the Alaska Qualified Assessor
r2 or R2	Rater 2 - The DRA expert reviewer

Data Analyses

A comprehensive set of analyses were conducted to determine the scoring accuracy of Alaska Assessors in scoring CWS, CLS, and I & O. Alaska Assessors were defined as the first rater (R1), while the expert rater from DRA was the second rater (R2). Several analyses were conducted, including descriptive statistics, Kappa statistics, and cross-tabulation tables. An anecdotal review of all test documents was also completed. For all analyses, items of a particular type were compared across each rater. The purpose of comparing R1 with R2 was to examine the degree to which ratings from the field (R1) matched with ratings from DRA's expert (R2).

Descriptive Statistics

The data presented in Table 1 provide the means (the average rating of student performance), standard deviations (the amount of variation around the mean), and count (for each item or task with complete data). Data are organized by item, by rater. The

variable n is the total number of items rated by the total number of students (e.g., the total number of possible ratings). The entries are grouped by rater to demonstrate the consistency between the raters for each aspect rated.

Table 1

Descriptive Statistics

Variable	N	Minimum	Maximum	Mean		Standard Deviation
				Statistic	Std. Error	
r1CWS1.56A1r	74	0	5	2.35	.258	2.223
r2CWS1.56A1r	74	0	5	2.30	.259	2.225
r1CWS1.56A2r	75	0	5	2.52	.258	2.238
r2CWS1.56A2r	75	0	5	2.43	.262	2.273
r1CLS1.56B1pct	75	0	100	60.70	4.891	42.360
r2CLS1.56B1pct	75	0	100	57.85	4.908	42.509
r1CLS1.56B1pts	75	0	25	14.27	1.262	10.926
r2CLS1.56B1pts	75	0	25	13.47	1.277	11.058
r1CLS1.56C1r	75	0	4	1.48	.187	1.622
r2CLS1.56C1r	75	0	4	1.45	.186	1.613
r1CLS1.56C2r	75	0	4	1.88	.201	1.740
r2CLS1.56C2r	75	0	4	1.85	.199	1.722
r1CLS1.56C3r	75	0	4	1.60	.201	1.740
r2CLS1.56C3r	75	0	4	1.57	.202	1.749
r1CLS1.56C4r	55	0	5	2.16	.253	1.873
r2CLS1.56C4r	55	0	5	2.09	.257	1.908
r1CLS1.56C5r	55	0	4	2.09	.201	1.494
r2CLS1.56C5r	55	0	4	2.04	.203	1.503
r1CWS1.56D1num	58	0	7	1.33	.234	1.781
r2CWS1.56D1num	58	0	7	1.34	.235	1.792
r1CWS1.56D1den	58	0	10	2.57	.370	2.817
r2CWS1.56D1den	58	0	10	2.75	.370	2.792
r1CWS1.56D1pct	58	0	100	26.71	4.533	34.522
r2CWS1.56D1pct	58	0	100	27.19	4.523	34.445
r1CWS1.56D1pts	58	0	10	2.72	.457	3.483
r2CWS1.56D1pts	58	0	10	2.79	.457	3.478
r1IO1.56D1	58	0	5	1.66	.210	1.596
r2IO1.56D1	58	0	5	1.62	.208	1.588
r1CWS1.56D2num	57	0	20	1.49	.407	3.071
r2CWS1.56D2num	57	0	6	1.12	.232	1.753
r1CWS1.56D2den	57	0	40	3.81	.807	6.096

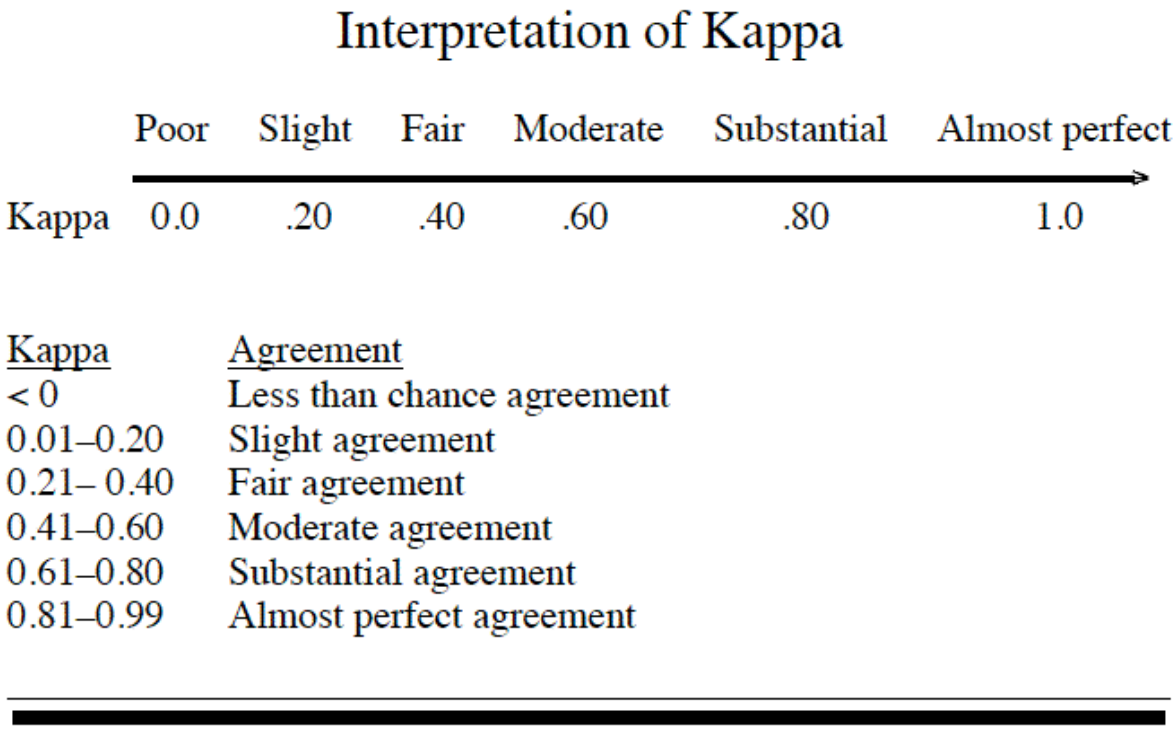
Variable	N	Minimum	Maximum	Mean		Standard Deviation
				Statistic	Std. Error	
r21.56D2den	57	0	16	3.35	.499	3.768
r1CWS1.56D2pct	57	0	100	19.59	3.810	28.767
r2CWS1.56D2pct	57	0	100	18.08	3.683	27.808
r1CWS1.56D2pts	57	0	10	2.05	.391	2.955
r2CWS1.56D2pts	57	0	10	1.91	.379	2.862
r1IO1.56D2	57	0	5	1.67	.216	1.629
r2IO1.56D2	57	0	5	1.67	.222	1.673
r11.56D3num	57	0	25	1.37	.475	3.589
r21.56D3num	57	0	8	.96	.220	1.658
r11.56D3den	57	0	43	3.82	.835	6.305
r21.56D3den	57	0	14	3.30	.474	3.576
r1CWS1.56D3pct	57	0	100	15.93	3.467	26.172
r2CWS1.56D3pct	57	0	100	15.28	3.389	25.587
r1CWS1.56D3pts	57	0	10	1.70	.357	2.699
r2CWS1.56D3pts	57	0	10	1.61	.348	2.631
r1IO1.56D3	57	0	5	1.67	.217	1.640
r2IO1.56D3	57	0	5	1.70	.222	1.679
Valid N (listwise)	51					

These descriptive statistics convey the mean rating by R1 compared to R2 as one aspect of agreement. Theoretically, both raters should have the same mean ratings. If the means are substantially different, however, then we could conclude that for at least some students R1 and R2 differed in their ratings. However, it is also important to note that the mean, as a statistic, is particularly sensitive to outliers. That is, if two raters agreed on all but one rating, but were very discrepant on that one rating, the means may differ by the same amount as if two raters consistently disagreed by a small amount.

Overall, the differences between means across all ratings was not severe. It is also apparent that the *n*-sizes drop by twenty after CWS item 1.56C3. This is the item where the participation rule for the Extended Levels of Support (ELOS) was met and subsequent items should not have been administered/scored for ELOS students. Those students should not have participated in the remainder of this standard administration, having moved to the ELOS assessment.

Kappa Statistics

The Kappa statistic is a measure of agreement between the ratings of R1 and R2. The following chart displays an interpretation guideline for the Kappa statistic, reprinted from J.R. Landis, et. al.[1] When a review includes two raters, and two raters only, the Kappa statistic should generally be interpreted as the primary measure of agreement. It is calculated from the cross-tabulation tables, and is interpreted similarly to a Pearson correlation coefficient.



Kappa	Agreement
< 0	Less than chance agreement
0.01–0.20	Slight agreement
0.21– 0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–0.99	Almost perfect agreement

The Kappa statistics are presented in Table 2 below. The Kappa statistics for Grade 5 range from .857 on CWS1.56D2num to .981 on CLS1.56C1r and CLS1.56C2r. These results reflect almost perfect agreement across all ratings in Grade 5. These results reflect higher levels of agreement than last year's Grade 8 and 10 analyses, in general.

It must be noted that there were some examples where the final score was the same even though the way in which the score was generated was different (the ^ differences matched the – differences in the word or sentence, so there was no net effect). In addition, 17.8 % of the overall scores were 0s. It is generally easier to determine that responses are incorrect than it is to generate matching scores for correct answers. These qualifications must be mentioned, as they both suggest that overall agreement may be slightly inflated. The slight inflation does not appear to be cause for concern, however.

[1] From "The Measurement of Observer Agreement for Categorical Data," by J.R. Landis, G.G. Koch, 1997, *Biometrics*, 33, p. 159-74.

Table 2

Kappa Statistics by Item

Comparison	Kappa
CWS1.56A1r	0.943
CWS1.56A2r	0.886
CLS1.56B1pct	0.878
CLS1.56B1pts	0.889
CLS1.56C1r	0.981
CLS1.56C2r	0.981
CLS1.56C3r	0.959
CLS1.56C4r	0.954
CLS1.56C5r	0.903
CWS1.56D1num	0.950
CWS1.56D1den	0.951
CWS1.56D1pct	0.950
CWS1.56D1pts	0.950
IO1.56D1	0.879
CWS1.56D2num	0.857
CWS1.56D2den	0.931
CWS1.56D2pct	0.861
CWS1.56D2pts	0.859
IO1.56D2	0.881
1.56D3num	0.970
1.56D3den	0.953
CWS1.56D3pct	0.913
CWS1.56D3pts	0.940
IO1.56D3	0.904

Graphs that convey the Kappa statistics visually can be found in *Appendix A, Kappa Plots*. These graphs convey a line of fit for each statistic, and also show the outliers that affected these results. One example of an impactful outlier can be seen with rating CWS1.56D2num. This rating received the lowest Kappa results (.857), as the Alaska Assessor scored a CWS item as a CLS item. The Assessor's numerator was severely inflated. One small error such as this resulted in a significant outlier that had a large impact upon the Kappa statistic, as well as the confidence interval around the line of fit.

Cross-tabulation Tables

The crosstabulation tables presented below in Tables 3-27 convey a representation of agreement between the ratings of R1 and R2 for Grade 5 across all items on the grade 5 writing assessment. The first rater is listed on the left, while the second rater is listed on the top. In terms of the coding for the headers, "r1CWS1.56A1r" means that it was rated by R1, that it was a CWS task, that the task number was 1.56A, that it was item 1, and that it was the raw score associated with that item. The numbers within each cell correspond to where individual ratings overlapped (i.e., agreements are on the diagonal), or differed. For example, in Table 3, R1 and R2 agreed upon 30/31 ratings of zero, with R1 awarding a 2 for one item that R2 scored as a zero. Consistent with the Kappa statistics presented above, the results reflect exceptional agreement and minimal cause for concern across all item types.

Table 3

r1CWS1.56A1r * r2CWS1.56A1r Crosstabulation								
Count								
		r2CWS1.56A1r						Total
		0	1	2	3	4	5	
r1CWS1.56A1r	0	30	0	0	0	0	0	30
	1	0	4	0	0	0	0	4
	2	1	0	3	0	0	0	4
	3	0	0	0	6	0	0	6
	4	0	0	0	2	4	0	6
	5	0	0	0	0	0	24	24
Total		31	4	3	8	4	24	74

Table 4

r1CWS1.56A2r * r2CWS1.56A2r Crosstabulation								
Count								
		r2CWS1.56A2r						Total
		0	1	2	3	4	5	
r1CWS1.56A2r	0	28	0	0	0	0	0	28
	1	0	4	1	0	0	0	5
	2	1	0	1	0	0	0	2
	3	0	1	0	6	0	0	7
	4	1	0	0	1	3	1	6
	5	0	0	0	0	0	27	27
Total		30	5	2	7	3	28	75

Table 5

r1CLS1.56B1pct * r2CLS1.56B1pct Crosstabulation																														
Count																														
		r2CLS1.56B1pct																												Total
		0	12	14	15	17	19	25	29	33	36	38	39	42	45	47	50	54	55	76	82	83	85	86	87	88	94	100	
r1CLS1.56B1pct	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18
	15	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	17	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	19	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	25	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	29	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	33	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	35	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	38	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	39	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
	45	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	50	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2
	54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
	55	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
	76	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
	77	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
	83	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
	85	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
	86	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
	87	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
	92	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	94	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
	100	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	30	32
Total		18	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	30	75

Table 6

r1CLS1.56B1pts * r2CLS1.56B1pts Crosstabulation								
Count								
		r2CLS1.56B1pts						Total
		0	5	10	15	20	25	
r1CLS1.56B1pts	0	21	0	0	0	0	0	21
	5	1	5	0	0	0	0	6
	10	1	0	6	0	0	0	7
	15	0	1	0	1	0	0	2
	20	0	0	1	0	6	0	7
	25	0	1	0	0	1	30	32
Total		23	7	7	1	7	30	75

Table 7

r1CLS1.56C1r * r2CLS1.56C1r Crosstabulation							
Count							
		r2CLS1.56C1r					Total
		0	1	2	3	4	
r1CLS1.56C1r	0	32	0	0	0	0	32
	1	0	13	0	0	0	13
	2	0	0	10	0	0	10
	3	0	1	0	1	0	2
	4	0	0	0	0	18	18
Total		32	14	10	1	18	75

Table 8

r1CLS1.56C2r * r2CLS1.56C2r Crosstabulation							
Count							
		r2CLS1.56C2r					Total
		0	1	2	3	4	
r1CLS1.56C2r	0	30	0	0	0	0	30
	1	0	2	0	0	0	2
	2	0	0	15	0	0	15
	3	0	0	0	3	0	3
	4	0	0	1	0	24	25
Total		30	2	16	3	24	75

Table 9

r1CLS1.56C3r * r2CLS1.56C3r Crosstabulation							
Count							
		r2CLS1.56C3r					Total
		0	1	2	3	4	
r1CLS1.56C3r	0	36	0	0	0	0	36
	1	1	2	0	0	0	3
	2	0	1	12	0	0	13
	3	0	0	0	1	0	1
	4	0	0	0	0	22	22
Total		37	3	12	1	22	75

Table 10

r1CLS1.56C4r * r2CLS1.56C4r Crosstabulation								
Count								
		r2CLS1.56C4r						Total
		0	1	2	3	4	5	
r1CLS1.56C4r	0	14	0	0	0	0	0	14
	1	1	11	0	0	0	0	12
	2	0	0	5	0	0	0	5
	3	1	0	0	10	0	0	11
	4	0	0	0	0	1	0	1
	5	0	0	0	0	0	12	12
Total		16	11	5	10	1	12	55

Table 11

r1CLS1.56C5r * r2CLS1.56C5r Crosstabulation							
Count							
		r2CLS1.56C5r					Total
		0	1	2	3	4	
r1CLS1.56C5r	0	12	0	0	0	0	12
	1	1	5	0	0	0	6
	2	0	1	16	1	0	18
	3	0	0	0	3	0	3
	4	0	0	1	0	15	16
Total		13	6	17	4	15	55

Table 12

r1CWS1.56D1num * r2CWS1.56D1num Crosstabulation									
Count									
		r2CWS1.56D1num							Total
		0	1	2	3	4	5	7	
r1CWS1.56D1num	0	30	0	0	0	0	0	0	30
	1	0	7	0	1	0	0	0	8
	2	0	1	6	0	0	0	0	7
	3	0	0	0	3	0	0	0	3
	4	0	0	0	0	6	0	0	6
	5	0	0	0	0	0	3	0	3
	7	0	0	0	0	0	0	1	1
Total		30	8	6	4	6	3	1	58

Table 13

r1CWS1.56D1den * r2CWS1.56D1den Crosstabulation											
Count											
		r2CWS1.56Dden									Total
		0	2	3	4	5	6	7	9	10	
r1CWS1.56Dden	0	25	0	1	0	1	0	0	0	0	27
	2	0	3	0	0	0	0	0	0	0	3
	4	0	0	0	10	0	0	0	0	0	10
	5	0	0	0	0	10	0	0	0	0	10
	6	0	0	0	0	0	1	0	0	0	1
	7	0	0	0	0	0	0	4	0	0	4
	9	0	0	0	0	0	0	0	1	0	1
	10	0	0	0	0	0	0	0	0	1	1
Total		25	3	1	10	11	1	4	1	1	57

Table 14

r1CWS1.56D1pct * r2CWS1.56D1pct Crosstabulation															
Count															
		r2CWS1.56D1pct													Total
		0	14	20	25	40	43	50	60	67	71	78	80	100	
r1CWS1.56D1pct	0	30	0	0	0	0	1	0	0	0	0	0	0	0	31
	14	0	1	0	0	0	0	0	0	0	0	0	0	0	1
	20	0	0	3	0	0	0	0	0	0	0	0	0	0	3
	25	0	0	0	2	0	0	0	0	0	0	0	0	0	2
	29	0	1	0	0	0	0	0	0	0	0	0	0	0	1
	40	0	0	0	0	1	0	0	0	0	0	0	0	0	1
	50	0	0	0	0	0	0	6	0	0	0	0	0	0	6
	60	0	0	0	0	0	0	0	3	0	0	0	0	0	3
	67	0	0	0	0	0	0	0	0	1	0	0	0	0	1
	71	0	0	0	0	0	0	0	0	0	1	0	0	0	1
	78	0	0	0	0	0	0	0	0	0	0	1	0	0	1
	80	0	0	0	0	0	0	0	0	0	0	0	2	0	2
	100	0	0	0	0	0	0	0	0	0	0	0	0	5	5
Total		30	2	3	2	1	1	6	3	1	1	1	2	5	58

Table 15

r1CWS1.56D1pts * r2CWS1.56D1pts Crosstabulation										
Count										
		r2CWS1.56D1pts								Total
		0	2	3	4	5	6	7	8	10	
r1CWS1.56D1pts	0	30	0	0	0	1	0	0	0	0	31
	2	0	4	0	0	0	0	0	0	0	4
	3	0	1	2	0	0	0	0	0	0	3
	4	0	0	0	1	0	0	0	0	0	1
	5	0	0	0	0	6	0	0	0	0	6
	6	0	0	0	0	0	3	0	0	0	3
	7	0	0	0	0	0	0	1	0	0	1
	8	0	0	0	0	0	0	0	4	0	4
	10	0	0	0	0	0	0	0	0	5	5
Total		30	5	2	1	7	3	1	4	5	58

Table 16

r1IO1.56D1 * r2IO1.56D1 Crosstabulation								
Count								
		r2IO1.56D1						Total
		0	1	2	3	4	5	
r1IO1.56D1	0	24	0	0	0	0	0	24
	1	0	4	1	0	0	0	5
	2	0	2	2	0	0	0	4
	3	0	0	0	17	1	0	18
	4	0	0	1	0	5	0	6
	5	0	0	0	0	0	1	1
Total		24	6	4	17	6	1	58

Table 17

r1CWS1.56D2num * r2CWS1.56D2num Crosstabulation									
Count									
		r2CWS1.56D2num							Total
		0	1	2	3	4	5	6	
r1CWS1.56D2num	0	33	1	0	0	0	0	0	34
	1	1	6	0	0	0	0	0	7
	2	0	0	2	0	0	0	0	2
	3	0	1	0	4	0	0	0	5
	4	0	0	0	1	2	0	0	3
	5	0	0	0	0	0	3	0	3
	6	0	0	0	0	0	0	2	2
	20	0	0	1	0	0	0	0	1
Total		34	8	3	5	2	3	2	57

Table 18

r1CWS1.56D2num * r2CWS1.56D2num Crosstabulation									
Count									
		r2CWS1.56D2num							Total
		0	1	2	3	4	5	6	
r1CWS1.56D2num	0	33	1	0	0	0	0	0	34
	1	1	6	0	0	0	0	0	7
	2	0	0	2	0	0	0	0	2
	3	0	1	0	4	0	0	0	5
	4	0	0	0	1	2	0	0	3
	5	0	0	0	0	0	3	0	3
	6	0	0	0	0	0	0	2	2
	20	0	0	1	0	0	0	0	1
Total		34	8	3	5	2	3	2	57

Table 19

r1CWS1.56D2den * r21.56D2den Crosstabulation														
Count														
		r21.56D2den												Total
		0	2	3	4	5	6	7	9	10	11	13	16	
r1CWS1.56D2den	0	24	1	0	0	0	0	0	0	0	0	0	0	25
	1	0	1	0	0	0	0	0	0	0	0	0	0	1
	2	0	3	0	0	0	0	0	0	0	0	0	0	3
	3	0	0	1	0	0	0	0	0	0	0	0	0	1
	4	0	0	0	5	0	0	0	0	0	0	0	0	5
	5	0	0	0	0	10	0	0	0	0	0	0	0	10
	6	0	0	0	0	0	2	0	0	0	0	0	0	2
	7	0	0	0	0	0	0	4	0	0	0	0	0	4
	9	0	0	0	0	0	0	0	2	0	0	0	0	2
	10	0	0	0	0	0	0	0	0	1	0	0	0	1
	13	0	0	0	0	0	0	0	0	0	0	1	0	1
	16	0	0	0	0	0	0	0	0	0	0	0	1	1
	40	0	0	0	0	0	0	0	0	0	1	0	0	1
Total		24	5	1	5	10	2	4	2	1	1	1	1	57

Table 20

r1CWS1.56D2pct * r2CWS1.56D2pct Crosstabulation																			
Count																			
		r2CWS1.56D2pct																	Total
		0	10	11	14	18	20	25	33	38	38	43	50	60	67	71	80	100	
r1CWS1.56D2pct	0	33	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	34
	14	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	20	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	3
	25	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2
	30	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	33	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
	38	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
	38	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
	43	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	2
	50	0	0	0	0	1	0	0	0	0	0	0	2	0	0	0	0	0	3
	60	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2
	67	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
	71	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
	80	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	2
	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2
Total		34	1	1	1	1	3	1	1	1	1	2	2	3	1	1	1	2	57

Table 21

r1CWS1.56D2pts * r2CWS1.56D2pts Crosstabulation											
Count											
		r2CWS1.56D2pts									Total
		0	1	2	3	4	5	6	7	8	10	
r1CWS1.56D2pts	0	33	0	1	0	0	0	0	0	0	0	34
	2	0	0	4	0	0	0	0	0	0	0	4
	3	1	1	0	1	0	0	0	0	0	0	3
	4	0	0	0	0	3	0	0	0	0	0	3
	5	0	0	1	0	0	4	0	0	0	0	5
	6	0	0	0	0	0	0	2	0	0	0	2
	7	0	0	0	0	0	0	0	1	0	0	1
	8	0	0	0	0	0	0	1	0	2	0	3
	10	0	0	0	0	0	0	0	0	0	2	2
Total		34	1	6	1	3	4	3	1	2	2	57

Table 22

r1IO1.56D2 * r2IO1.56D2 Crosstabulation								
Count								
		r2IO1.56D2						Total
		0	1	2	3	4	5	
r1IO1.56D2	0	24	0	0	0	0	0	24
	1	0	3	0	0	0	0	3
	2	0	2	6	1	0	0	9
	3	0	0	0	10	0	1	11
	4	0	0	0	1	8	0	9
	5	0	0	0	0	0	1	1
Total		24	5	6	12	8	2	57

Table 23

r11.56D3num * r21.56D3num Crosstabulation										
Count										
		r21.56D3num								Total
		0	1	2	3	4	5	6	8	
r11.56D3num	0	35	0	0	0	0	0	0	0	35
	1	0	7	0	0	0	0	0	0	7
	2	0	0	7	0	0	0	0	0	7
	3	0	0	0	3	0	0	0	0	3
	4	0	0	0	0	1	0	0	0	1
	5	0	0	0	0	0	1	0	0	1
	6	0	0	0	0	0	0	1	0	1
	8	0	0	0	0	0	0	0	1	1
	25	0	0	1	0	0	0	0	0	1
Total		35	7	8	3	1	1	1	1	57

Table 24

r11.56D3den * r21.56D3den Crosstabulation												
Count												
		r21.56D3den										Total
		0	2	3	4	5	6	7	8	11	14	
r11.56D3den	0	25	1	0	0	0	0	0	0	0	0	26
	2	0	1	0	0	0	0	0	0	0	0	1
	3	0	0	2	0	0	0	0	0	0	0	2
	4	0	0	0	5	0	0	0	0	0	0	5
	5	0	0	0	0	11	0	0	0	0	0	11
	6	0	0	0	0	0	2	0	0	0	0	2
	7	0	0	0	0	0	0	4	0	0	0	4
	8	0	0	0	0	0	0	0	2	0	0	2
	11	0	0	0	0	0	0	0	0	2	0	2
	14	0	0	0	0	0	0	0	0	0	1	1
	43	0	0	0	0	0	0	0	0	1	0	1
Total		25	2	2	5	11	2	4	2	3	1	57

Table 25

r1CWS1.56D3pct * r2CWS1.56D3pct Crosstabulation															
Count															
		r2CWS1.56D3pct													Total
		0	13	14	18	20	29	33	36	50	60	66	71	100	
r1CWS1.56D3pct	0	35	0	0	0	0	0	0	0	0	0	0	0	0	35
	14	0	0	2	0	0	1	0	0	0	0	0	0	0	3
	18	0	0	0	1	0	0	0	0	0	0	0	0	0	1
	20	0	0	0	0	5	0	0	0	0	0	0	0	0	5
	25	0	1	0	0	0	0	0	0	0	0	0	0	0	1
	29	0	0	0	0	0	1	0	0	0	0	0	0	0	1
	33	0	0	0	0	0	0	1	0	0	0	0	0	0	1
	36	0	0	0	0	0	0	0	1	0	0	0	0	0	1
	50	0	0	0	0	0	0	0	0	1	0	0	0	0	1
	58	0	0	0	1	0	0	0	0	0	0	0	0	0	1
	60	0	0	0	0	0	0	0	0	0	3	0	0	0	3
	66	0	0	0	0	0	0	0	0	0	0	1	0	0	1
	71	0	0	0	0	0	0	0	0	0	0	0	1	0	1
	100	0	0	0	0	0	0	0	0	0	0	0	0	2	2
Total		35	1	2	2	5	2	1	1	1	3	1	1	2	57

Table 26

r1CWS1.56D3pts * r2CWS1.56D3pts Crosstabulation											
Count											
		r2CWS1.56D3pts									Total
		0	2	3	4	5	6	7	8	10	
r1CWS1.56D3pts	0	35	0	0	0	0	0	0	0	0	35
	2	0	8	0	0	0	0	0	0	0	8
	3	0	1	2	0	0	0	0	0	0	3
	4	0	0	0	2	0	0	0	0	0	2
	5	0	0	0	0	1	0	0	0	0	1
	6	0	1	0	0	0	3	0	0	0	4
	7	0	0	0	0	0	0	1	0	0	1
	8	0	0	0	0	0	0	0	1	0	1
	10	0	0	0	0	0	0	0	0	2	2
Total		35	10	2	2	1	3	1	1	2	57

Table 27

r1IO1.56D3 * r2IO1.56D3 Crosstabulation								
Count								
		r2IO1.56D3						Total
		0	1	2	3	4	5	
r1IO1.56D3	0	24	0	0	0	0	0	24
	1	0	3	1	0	0	0	4
	2	0	1	5	1	0	0	7
	3	0	0	0	12	0	0	12
	4	0	0	0	0	8	1	9
	5	0	0	0	0	0	1	1
Total		24	4	6	13	8	2	57

Summary & Implications for Future Training

The results from this year's writing scoring accuracy project demonstrate high levels of rater accuracy for Alaska Qualified Assessors.

Descriptive statistics reveal large variations among the student population, with standard deviations approaching or exceeding mean scores in many cases. The descriptive statistics also demonstrate consistent ratings between the two raters (R1 = Alaska Assessor/ R2 = Expert Rater). The ratings are close together, with means and standard deviations reflecting largely consistent results.

The Kappa statistics for Grade 5 range from .857 on CWS1.56D2num to .981 on CLS1.56C1r and CLS1.56C2r. These results reflect almost perfect agreement across all ratings in Grade 5 with higher levels of agreement compared to last year's Grade 8 and 10 analyses; this finding suggests either that Grade 5 responses are very easy to score, or that the resources that EED and DRA have dedicated to improving the writing scoring process appear to have been impactful. Raters appear to have good command of scoring CLS, which was the primary focus of this year's analyses. Refresher training on scoring CWS does not appear to be required. While CWS rater agreement remained lower than other areas, the results of the analyses presented here are very positive and suggest quite strong agreement.

Cross-tabulation tables demonstrated a high degree of scoring accuracy, with most discrepancies limited to one unit above or below consensus ratings.

Based on these results, the writing scoring rules appear to require no modifications. However, based upon the document review summary below, the field could benefit from additional guidance and training in specific areas of scoring writing.

Note: The Student Response column on the scoring protocol for the writing test appears to be unnecessary, as the student's work product should be in the student materials. It is recommended that this column be removed.

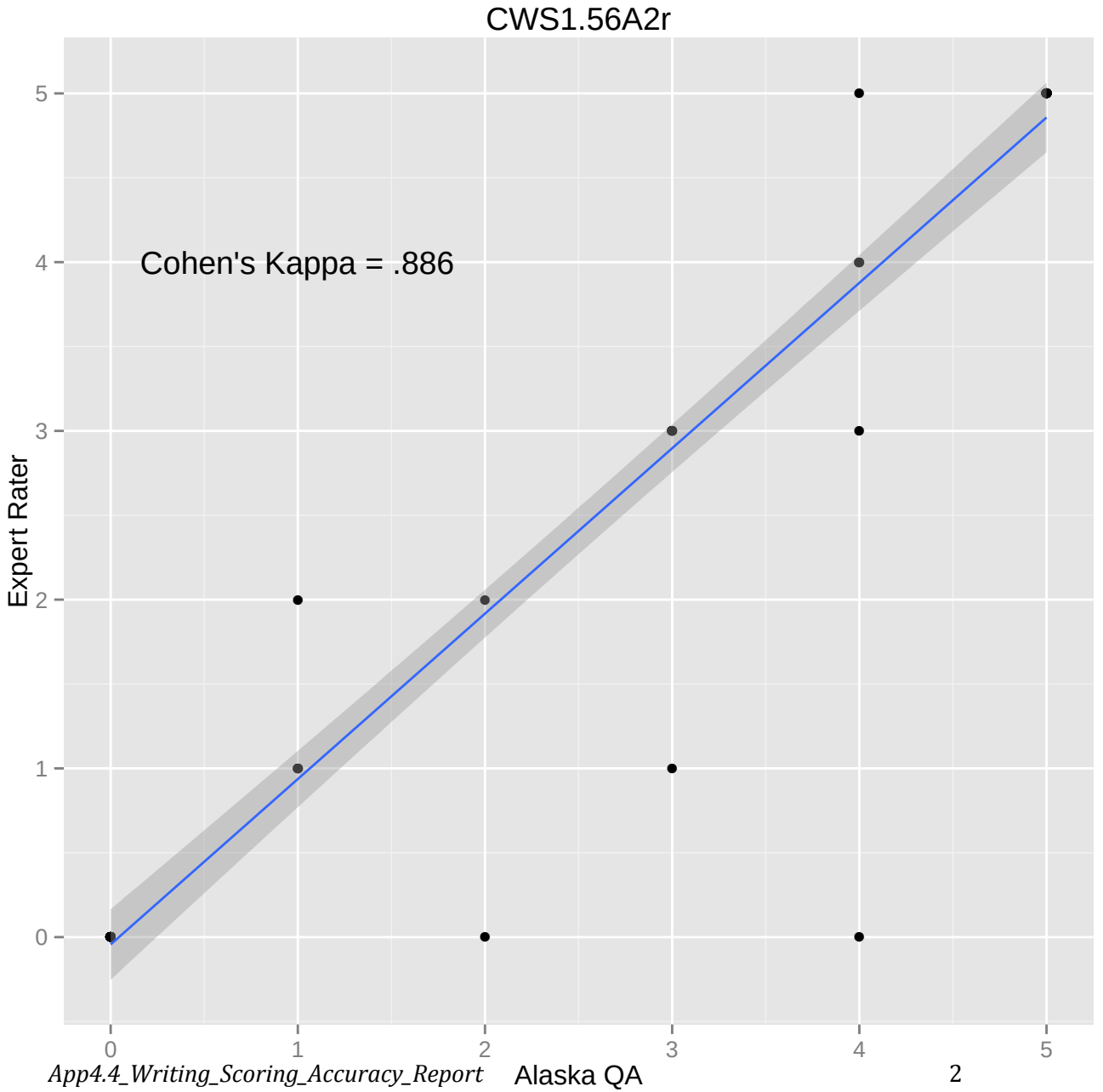
Document Review

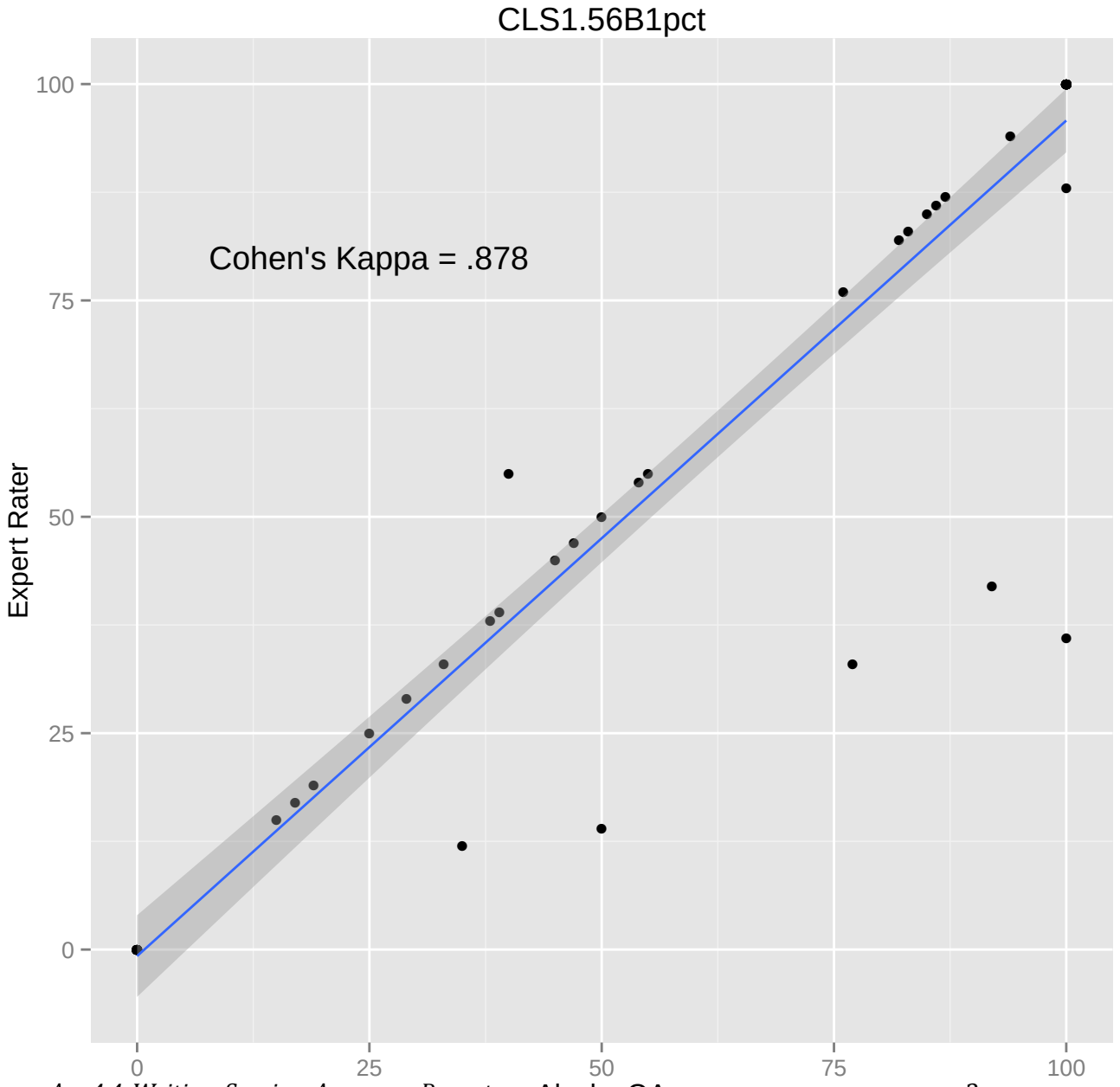
There were several patterns noted by DRA's expert reviewer during the score-behind process. The expert reviewer looked at the documentation present in all submitted scoring protocols and student materials to review accuracy of representation as well as self-consistency. Several areas of potential improvement were noted during this review as patterns emerged. The following topics are recommended for discussion during the next writing training for QAs and QTs:

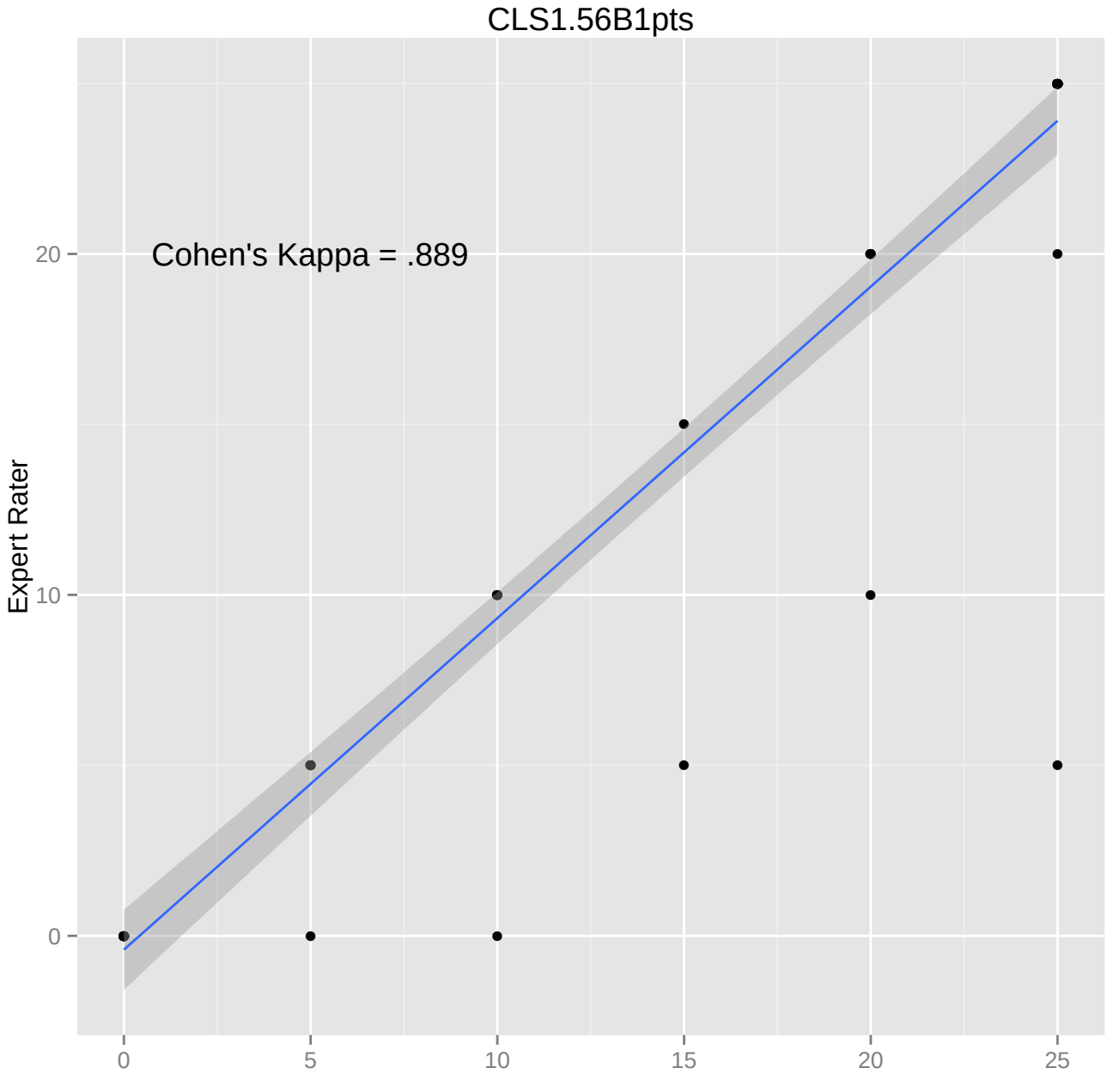
- One student wrote a first name twice; the Assessor scored the first name twice instead of scoring the actual last name as 0 points out of however many points were possible. It is recommended that this be discussed during training.
- Some Assessors continue to give credit for capital letters within sentences (F, N, P, L, etc.), even though credit should not be awarded unless the student uses all capitals. This needs to be re-emphasized in training.

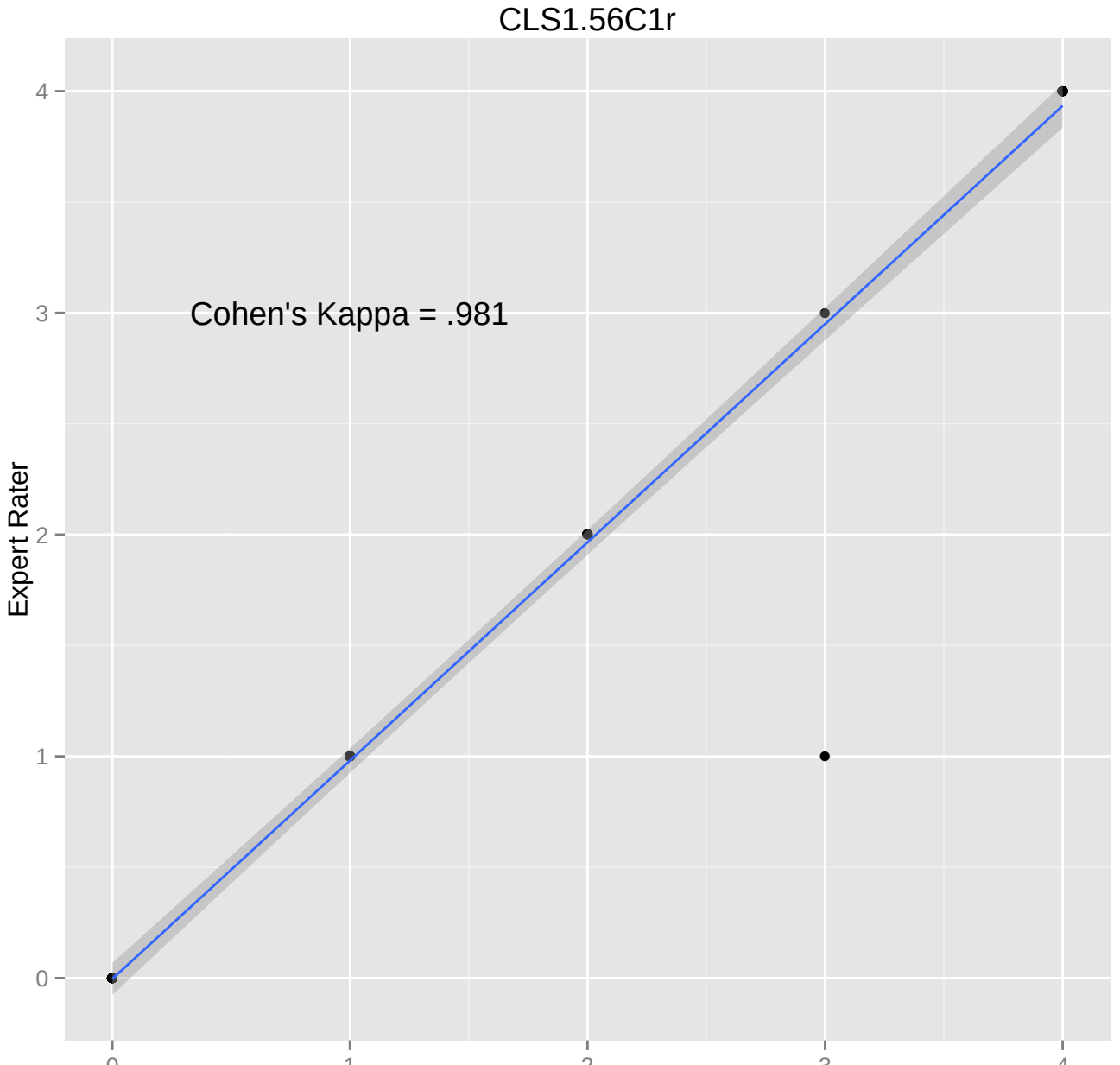
- Some Assessors appear to still be confused about scoring letter sequences and not the accuracy of each individual letter. For example, if the word were "This" and the student spells "^T-n-i^s^" it cannot be a 4/5, but a 3/5.
- One student insisted on using vertical lines to separate each word; the Assessor treated them as spaces and scored them as if they were simply demarcating appropriate spacing.
- Assessors appear to be generally less rigorous (too lenient) when grading the students when they were spelling their own names, possibly due to the teacher's familiarity with the student's name. The same CLS rules apply in the name-spelling tasks, however.
- One Assessor used the CLS scoring rules to score a CWS task.

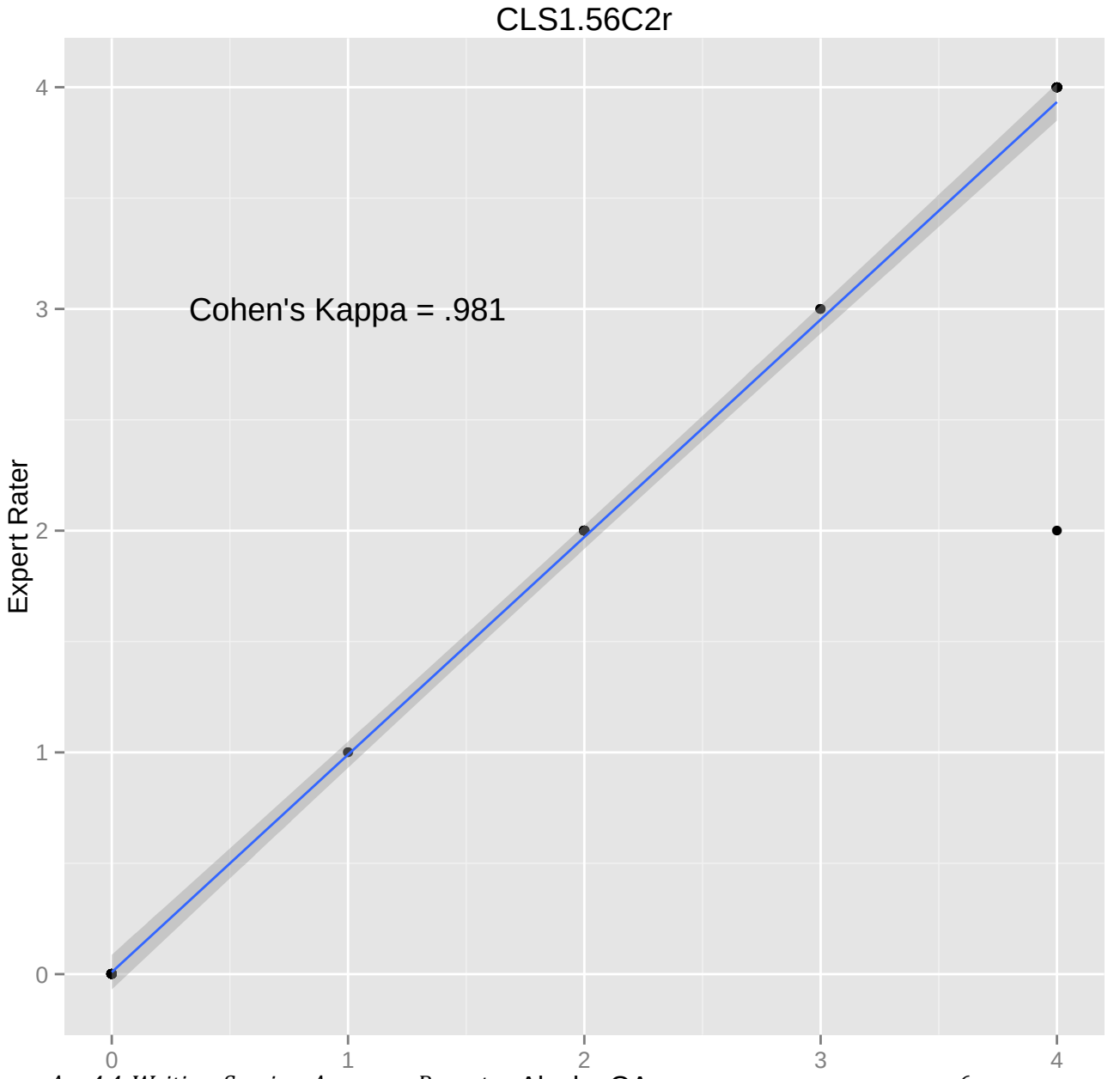

CWS1.56A1r
Cohen's Kappa = .943
Expert Rater
0
1
2
3
4
5
0
1
2
3
4
5

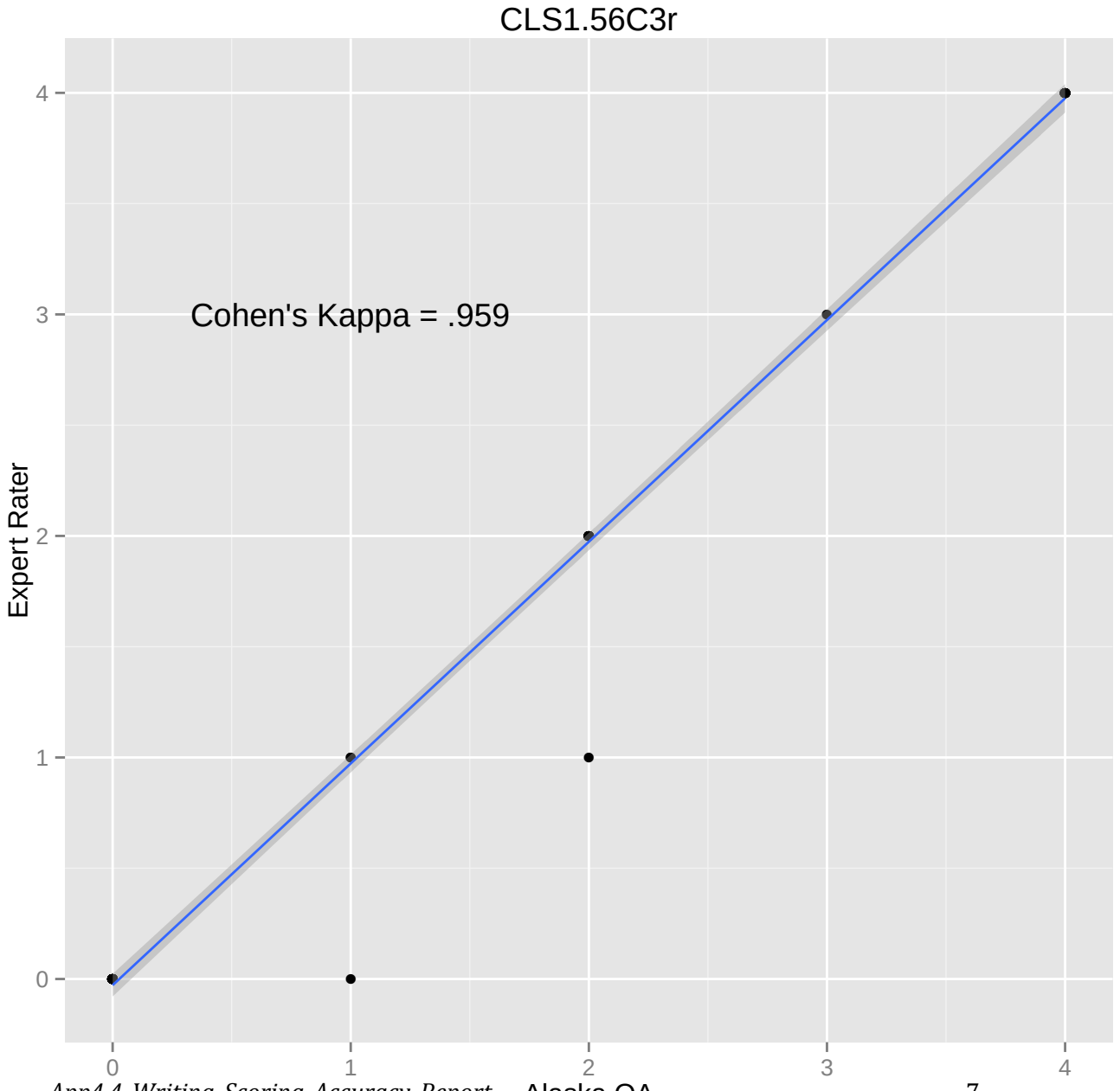

CWS1.56A2r
Cohen's Kappa = .886
Expert Rater
Alaska QA
0
1
2
3
4
5

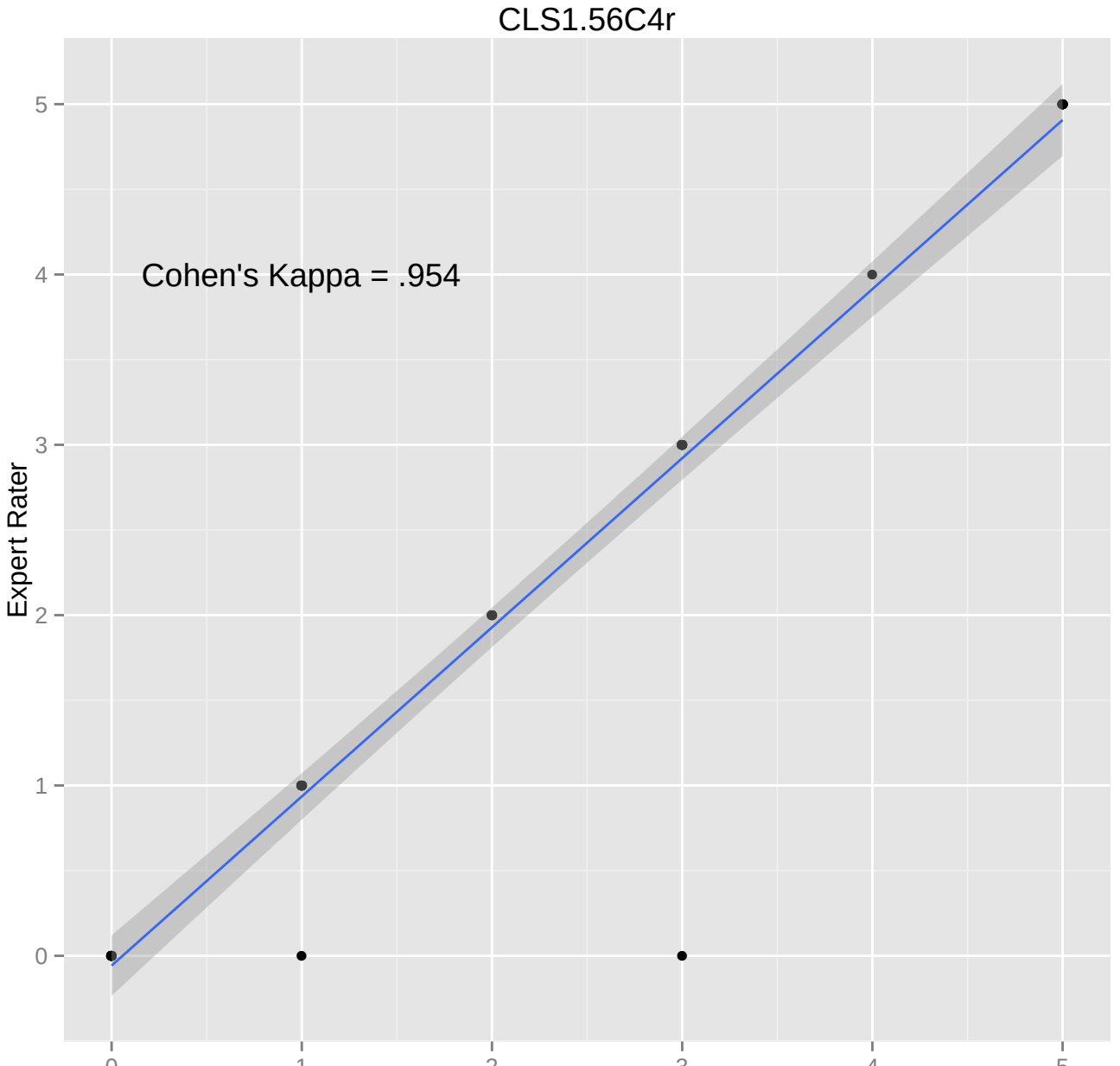

CLS1.56B1pct
Cohen's Kappa = .878
Expert Rater
100
75
50
25
0
0
25
50
75
100

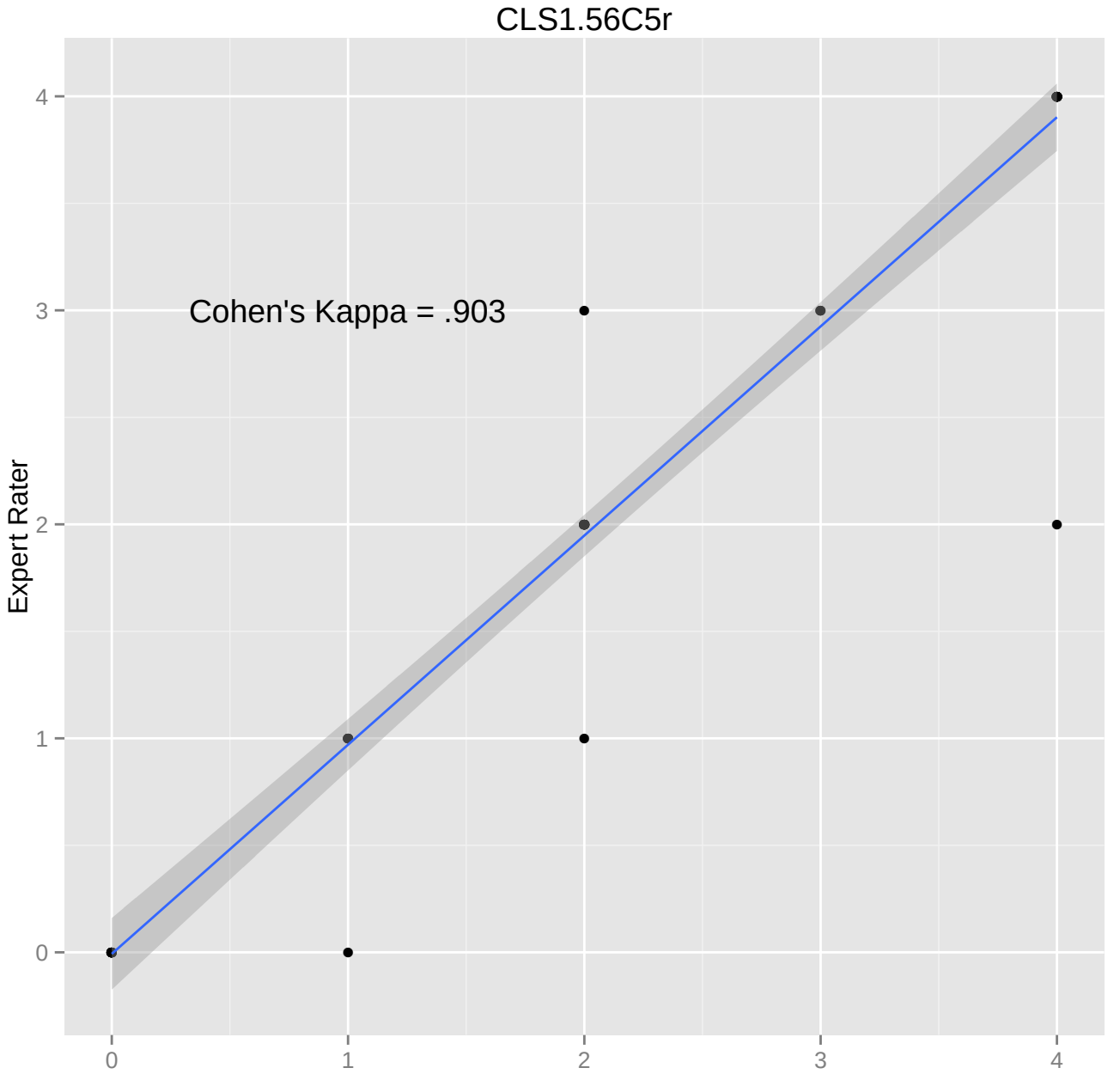

CLS1.56B1pts
Cohen's Kappa = .889
Expert Rater
20
10
0
0
5
10
15
20
25
Alaska QA

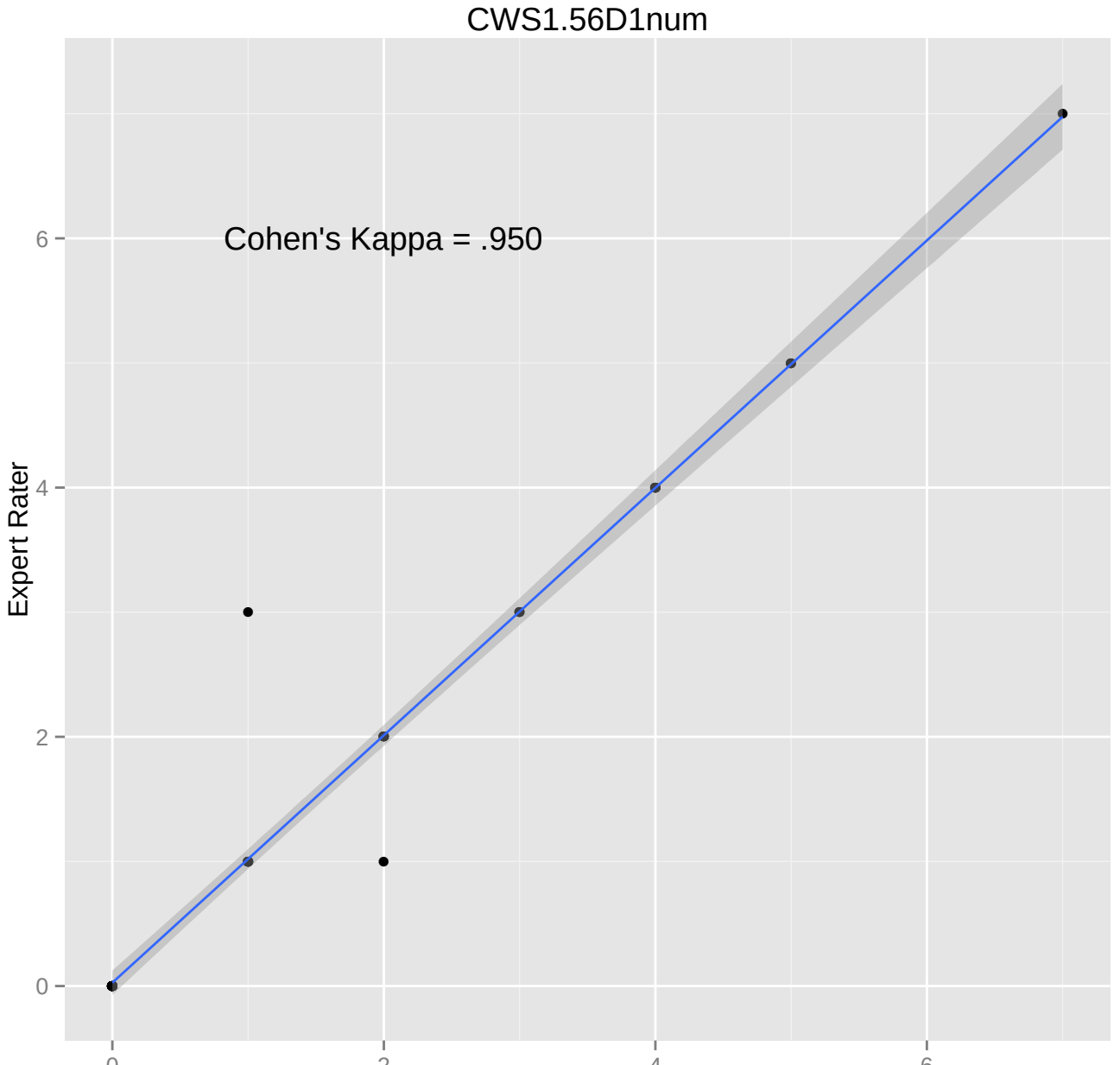

CLS1.56C1r
Cohen's Kappa = .981
Expert Rater
0
1
2
3
4
0
1
2
3
4

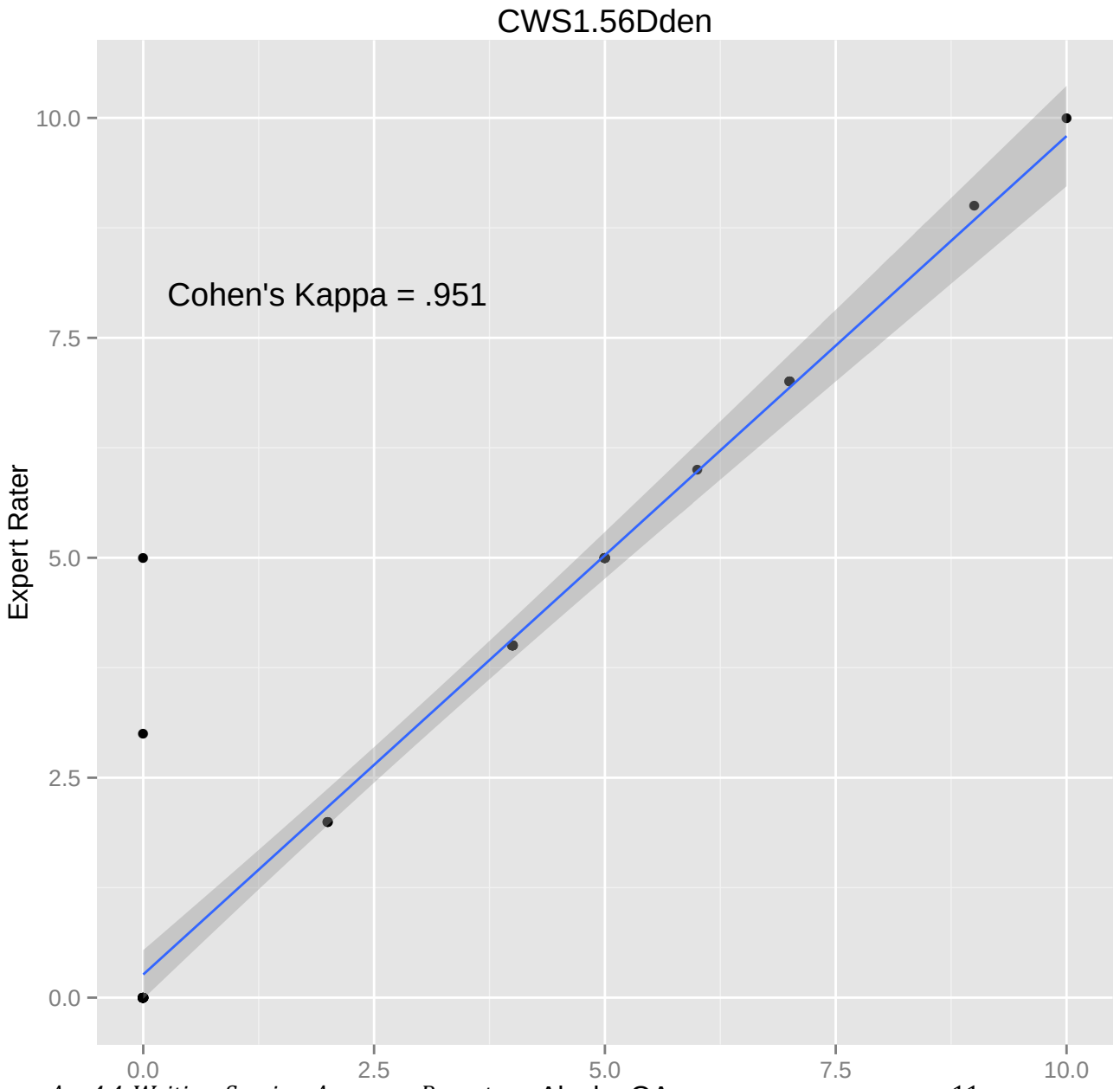

CLS1.56C2r
Cohen's Kappa = .981
Expert Rater
0
1
2
3
4

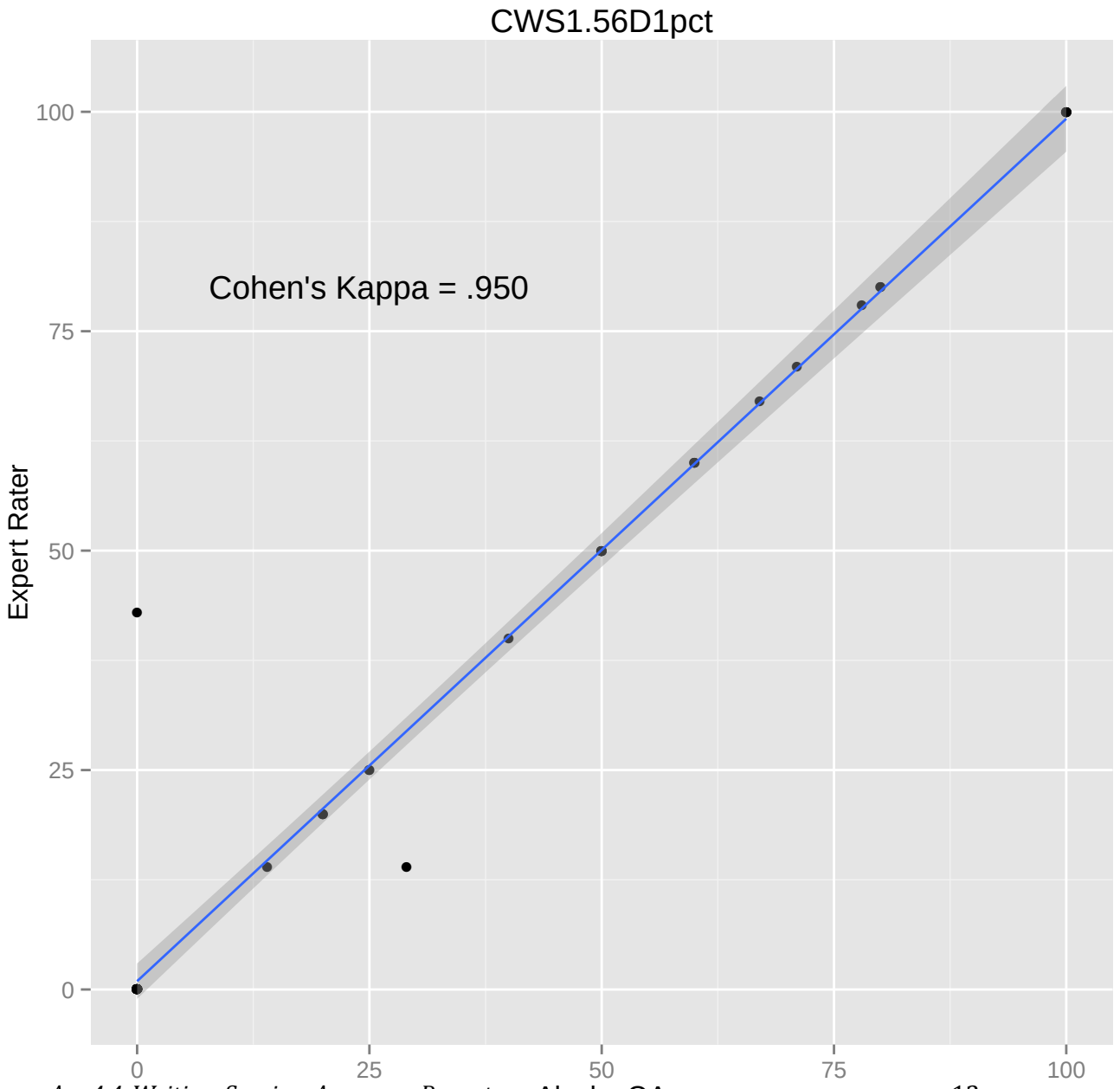

CLS1.56C3r
Cohen's Kappa = .959
Expert Rater
0
1
2
3
4
0
1
2
3
4

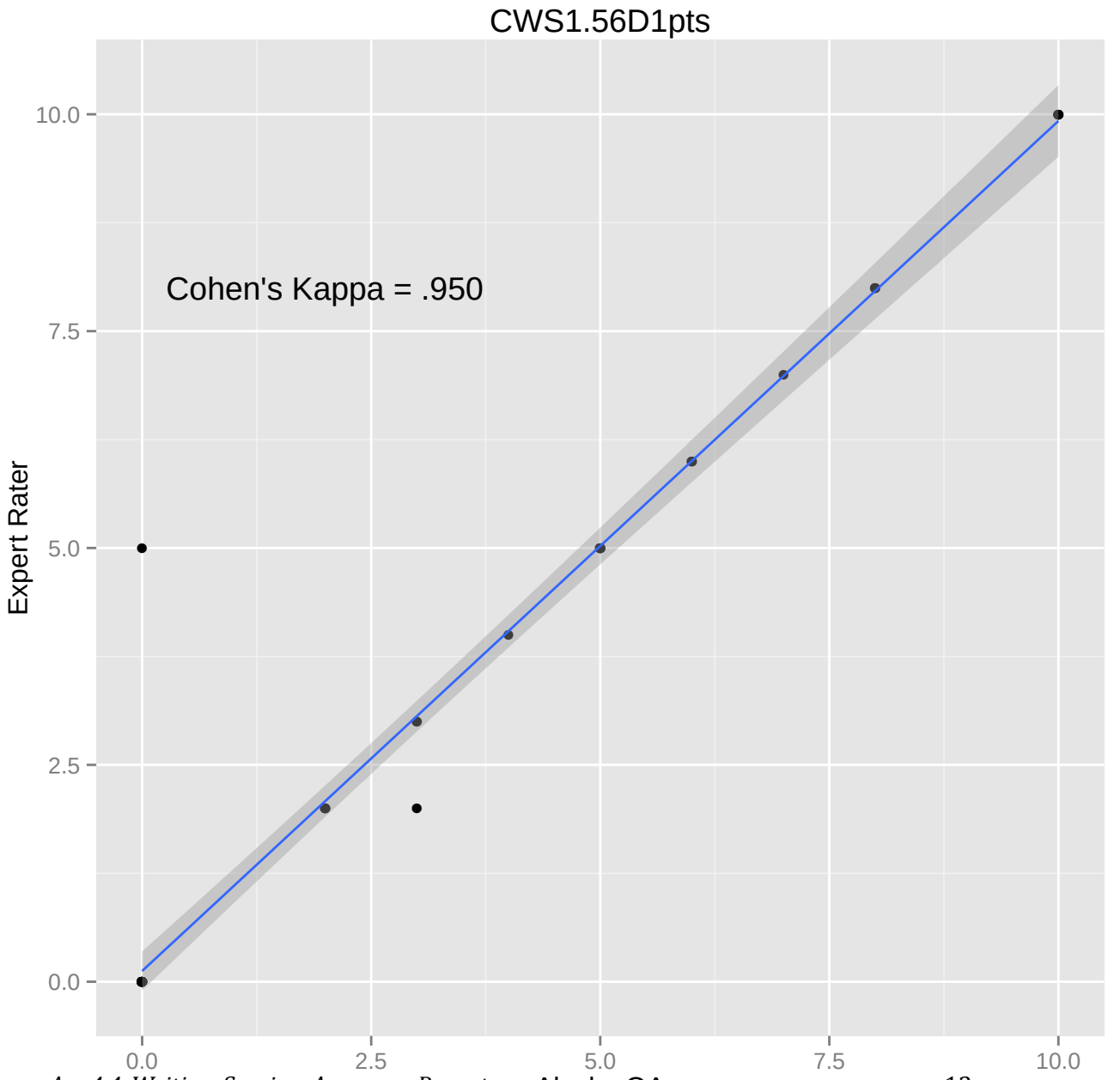

CLS1.56C4r
Cohen's Kappa = .954
Expert Rater
0
1
2
3
4
5

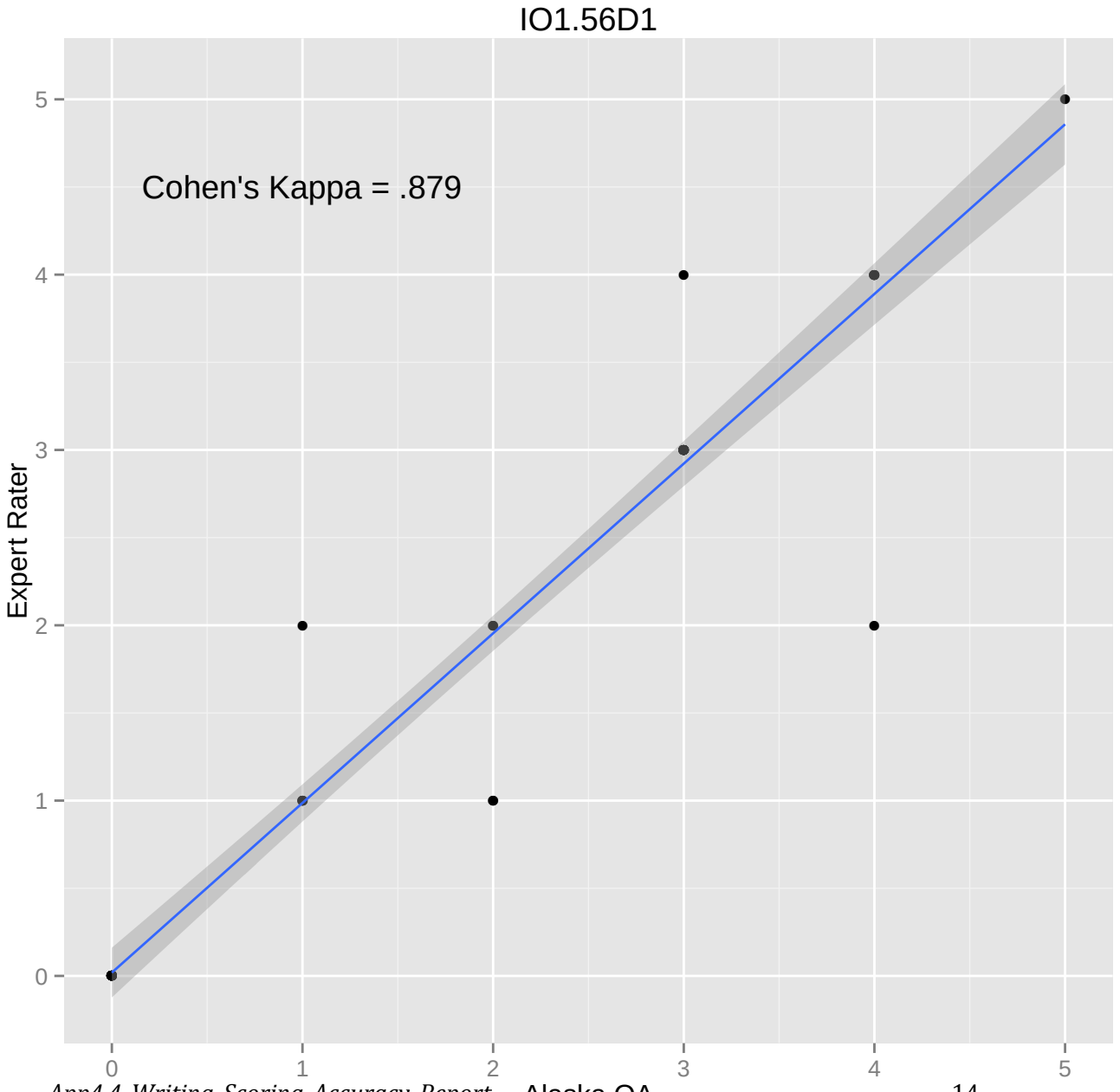

CLS1.56C5r
Cohen's Kappa = .903
Expert Rater
Alaska QA
0
1
2
3
4

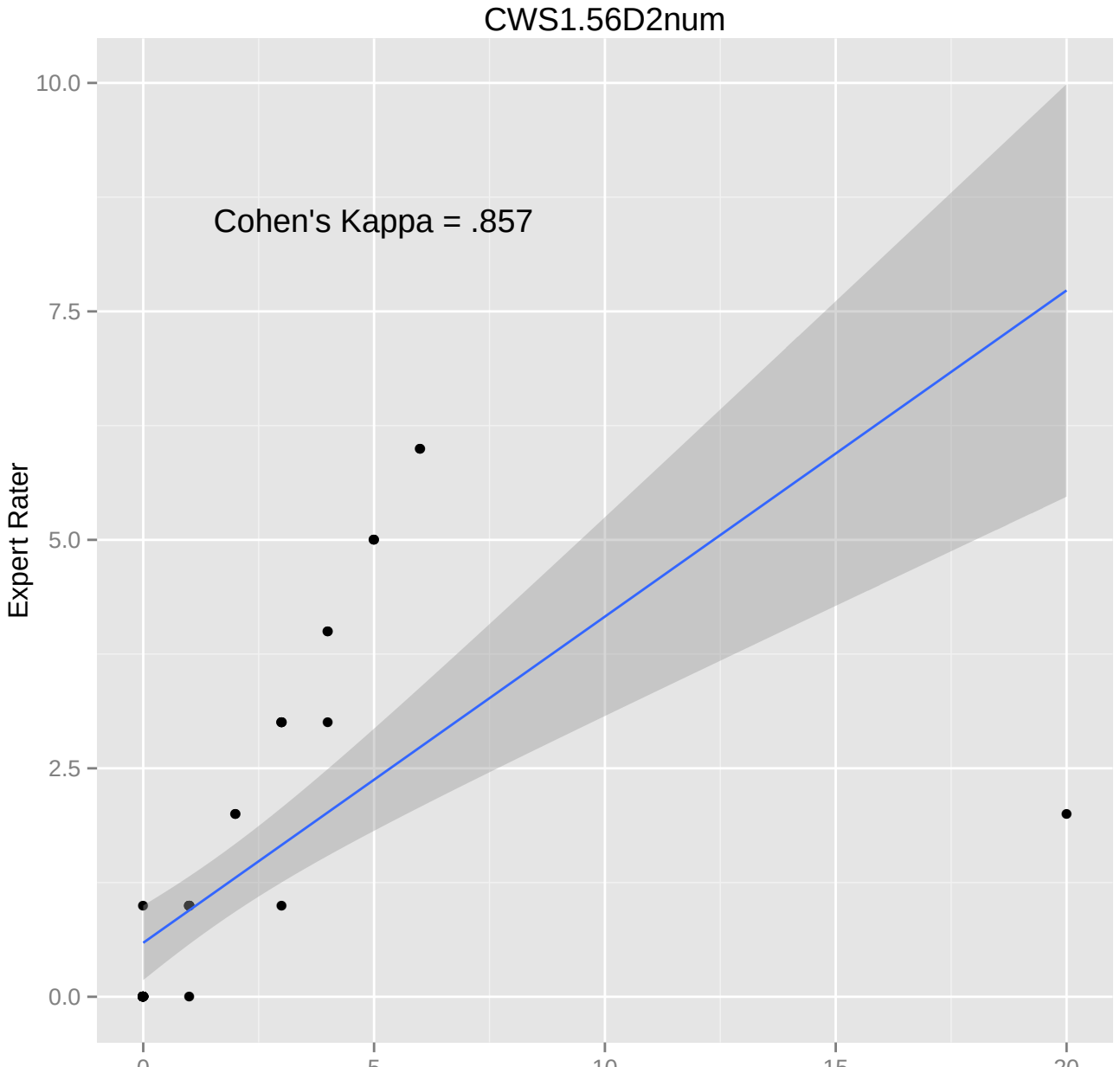

CWS1.56D1num
Cohen's Kappa = .950
Expert Rater
0
2
4
6
0
2
4
6

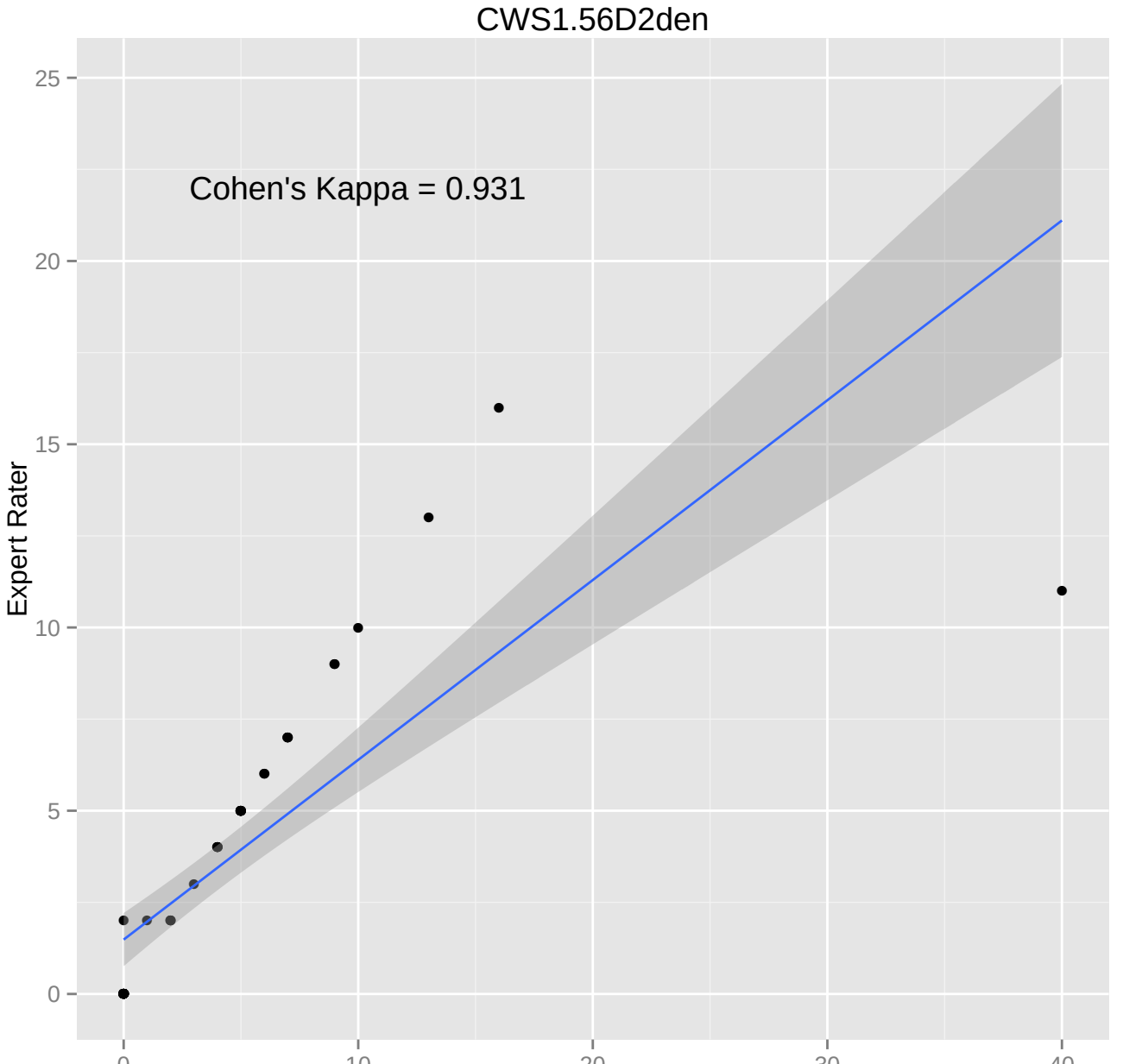

CWS1.56Dden
Cohen's Kappa = .951
Expert Rater
Alaska QA
10.0
7.5
5.0
2.5
0.0
0.0
2.5
5.0
7.5
10.0

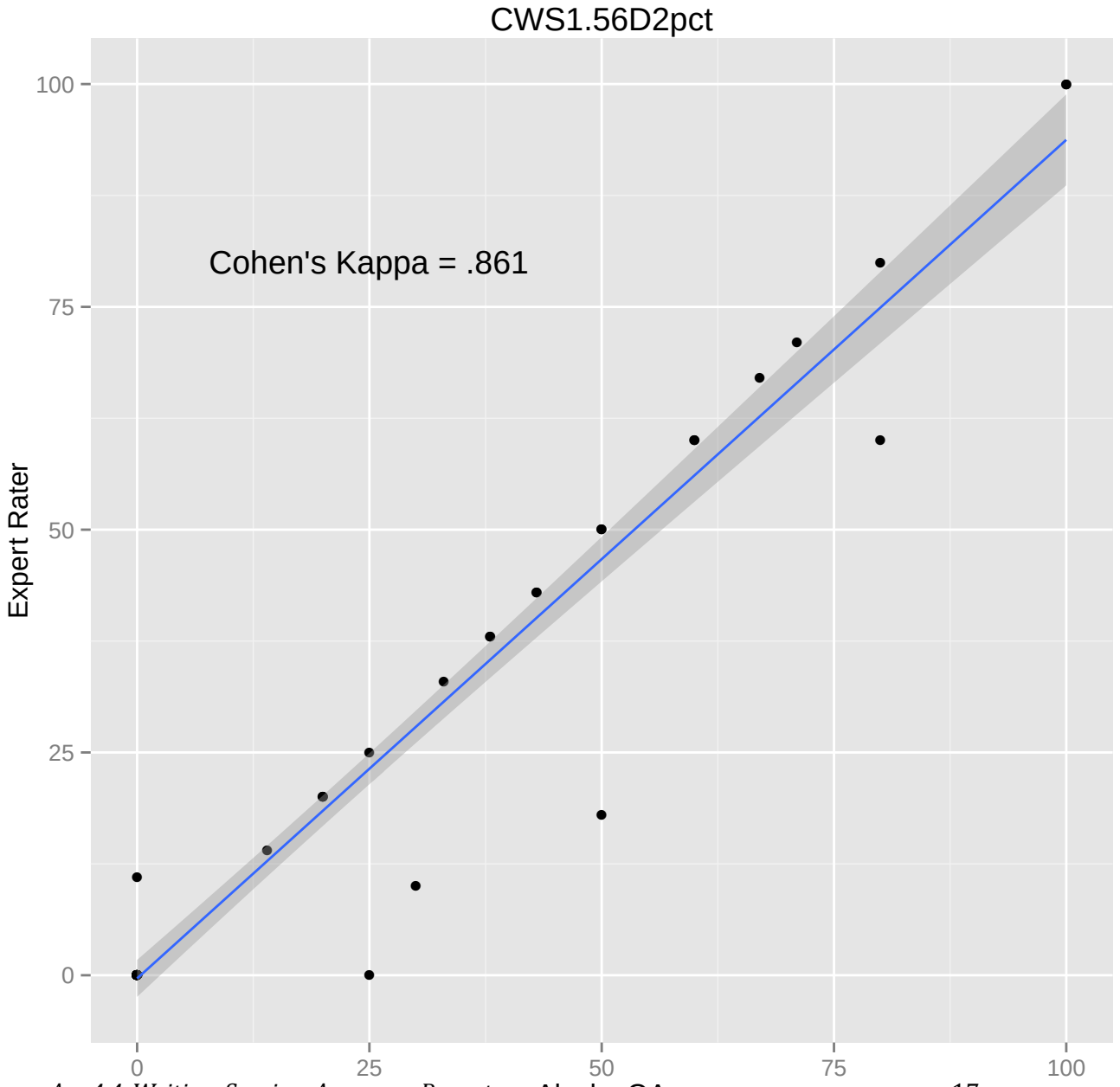

CWS1.56D1pct
Cohen's Kappa = .950
Expert Rater
100
75
50
25
0
0
25
50
75
100
Alaska QA

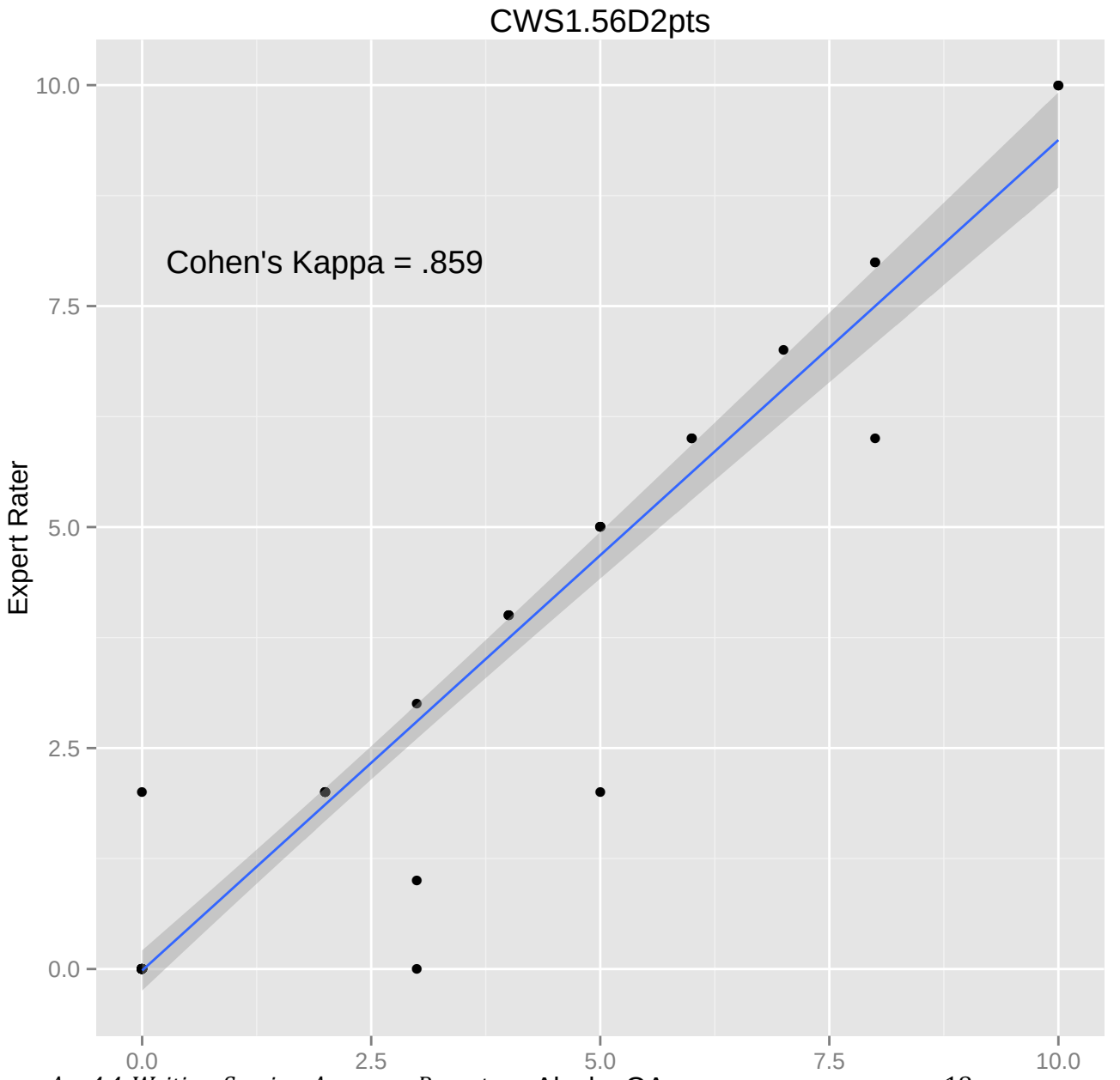

CWS1.56D1pts
Cohen's Kappa = .950
Expert Rater
Alaska QA
10.0
7.5
5.0
2.5
0.0
0.0
2.5
5.0
7.5
10.0

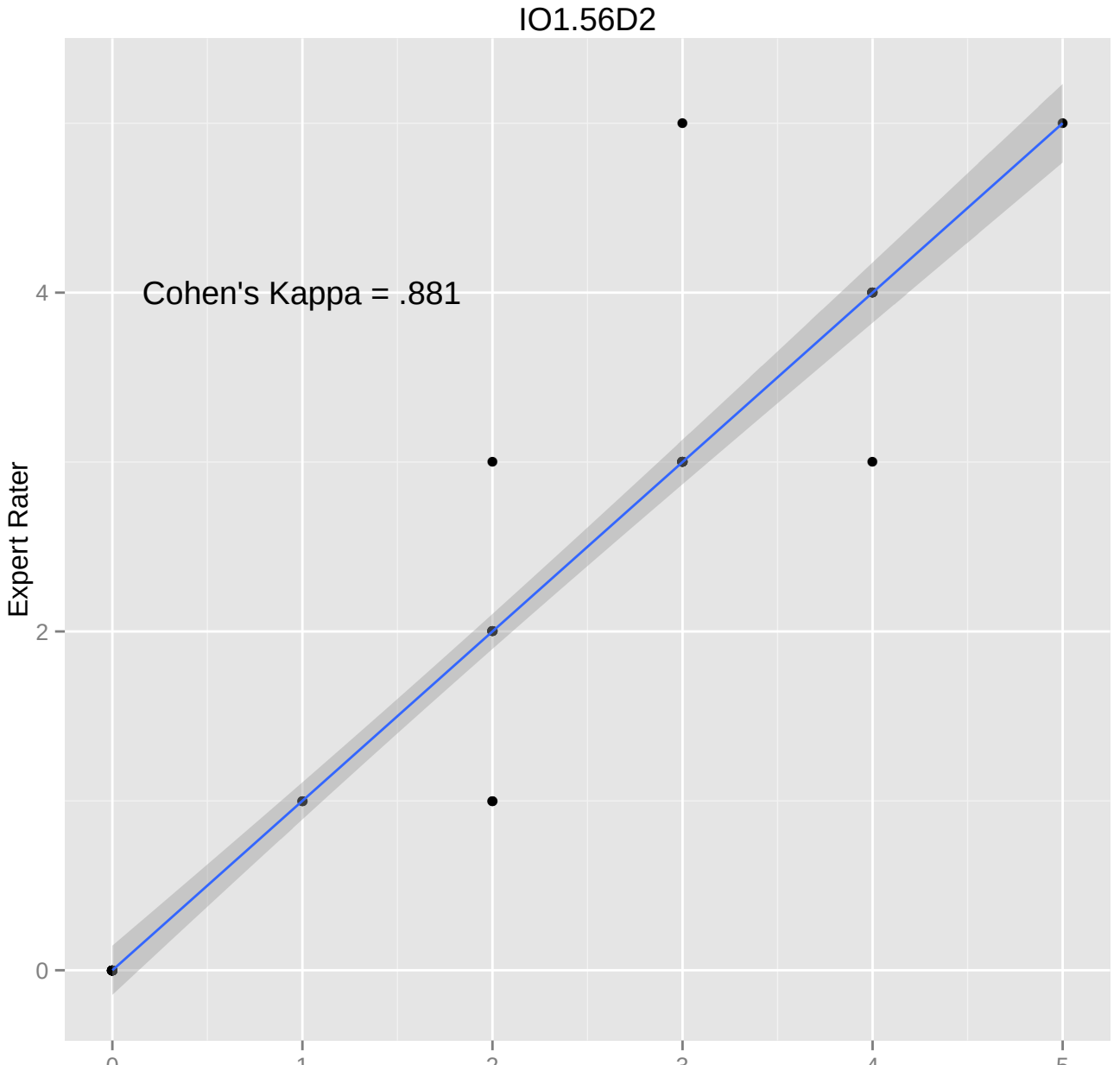

IO1.56D1
Cohen's Kappa = .879
Expert Rater
Alaska QA
0
1
2
3
4
5

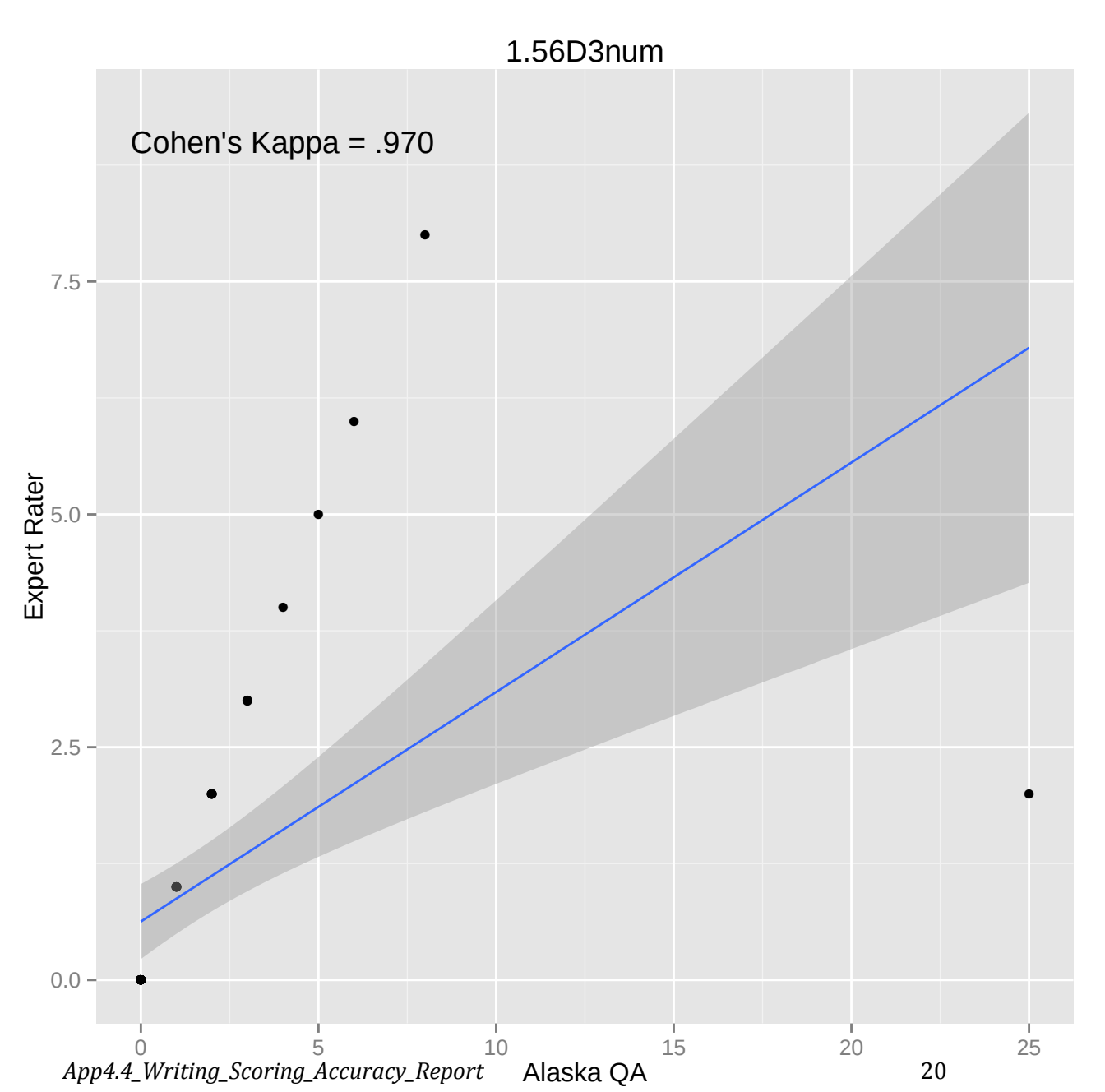

CWS1.56D2num
Cohen's Kappa = .857
Expert Rater
10.0
7.5
5.0
2.5
0.0
0
5
10
15
20
Alaska QA

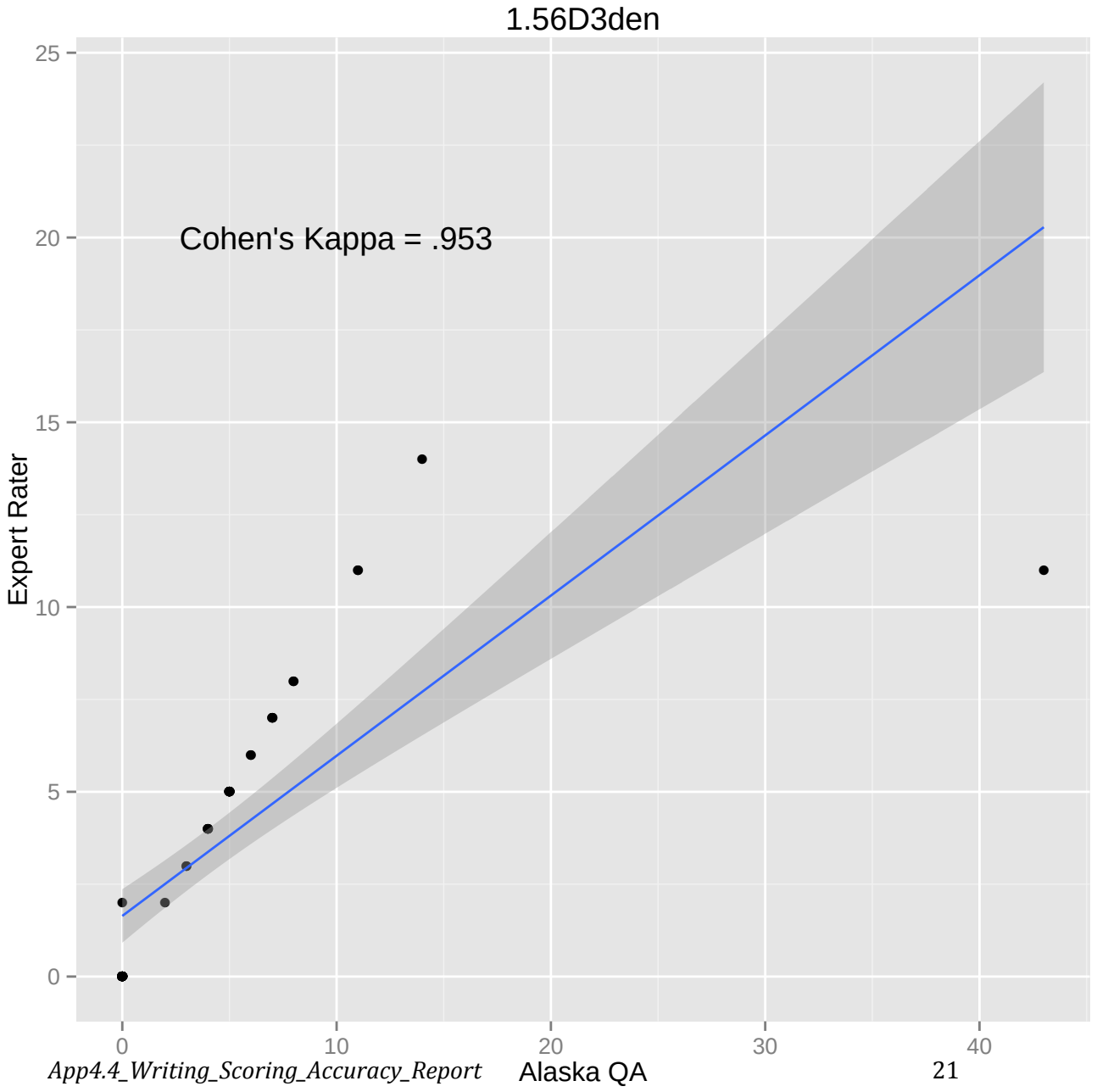

CWS1.56D2den
Cohen's Kappa = 0.931
Expert Rater
25
20
15
10
5
0
0
10
20
30
40

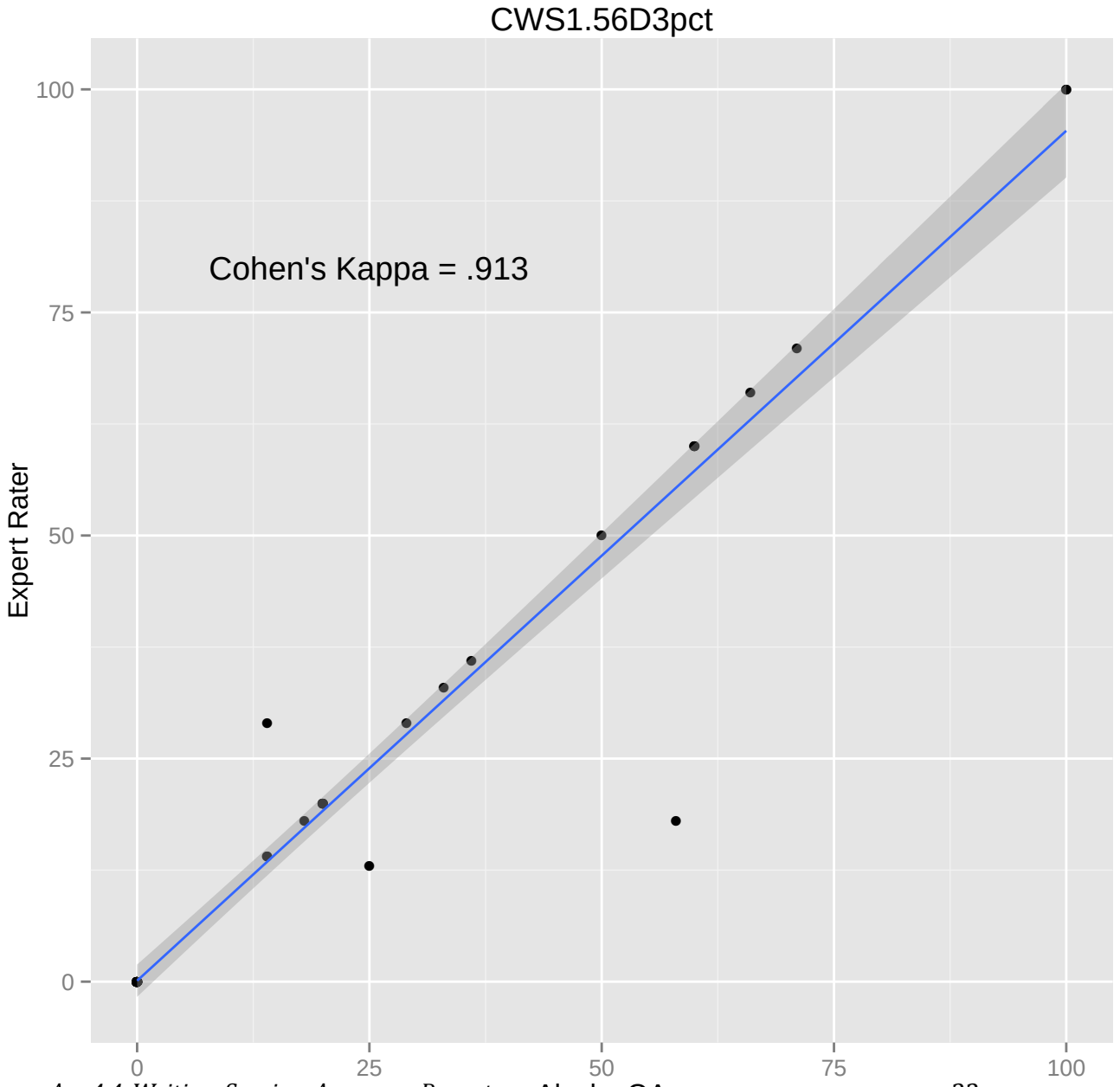

CWS1.56D2pct
Cohen's Kappa = .861
Expert Rater
100
75
50
25
0
0
25
50
75
100
Alaska QA

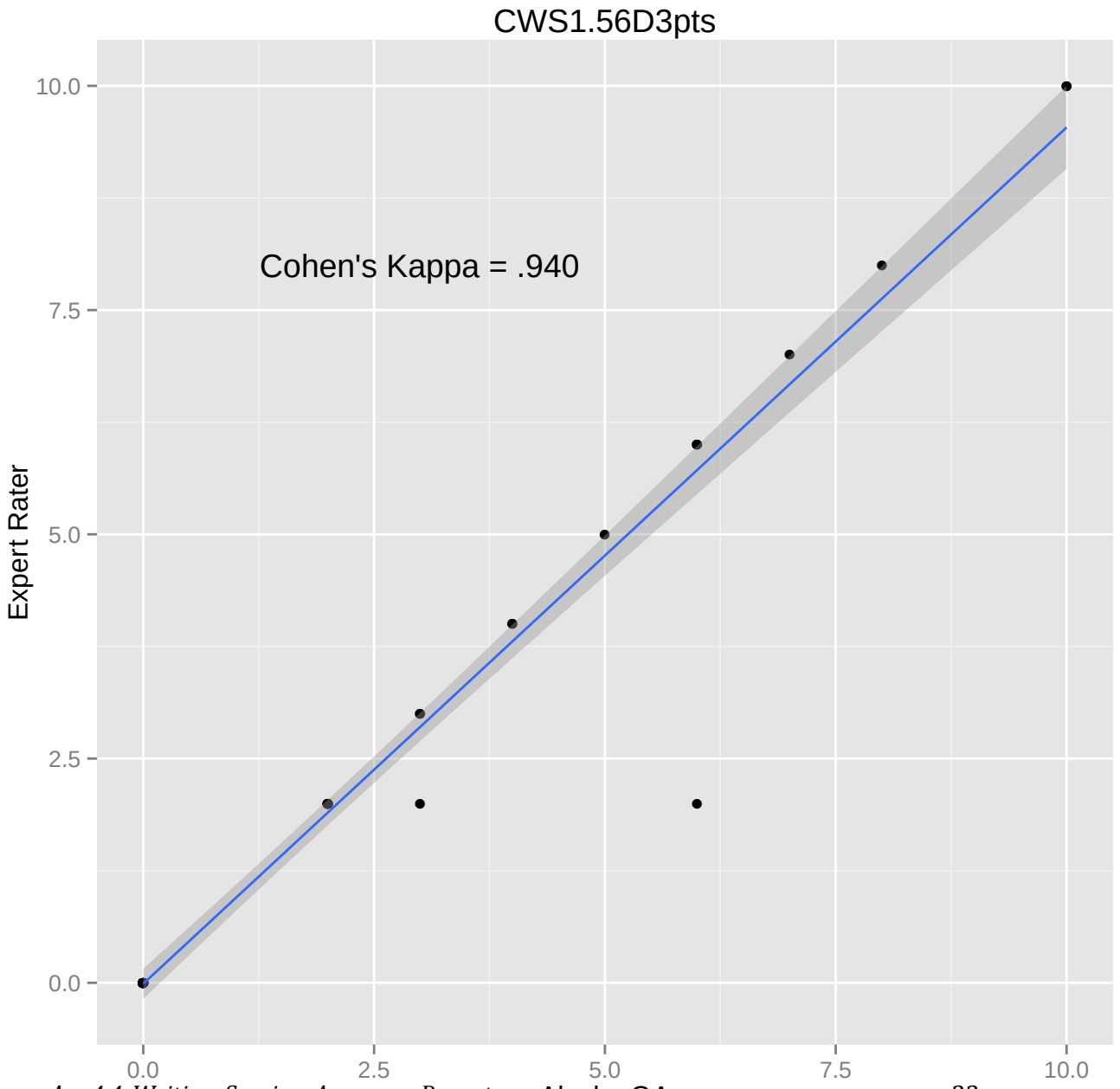

CWS1.56D2pts
Cohen's Kappa = .859
Expert Rater
10.0
7.5
5.0
2.5
0.0
0.0
2.5
5.0
7.5
10.0
Alaska QA

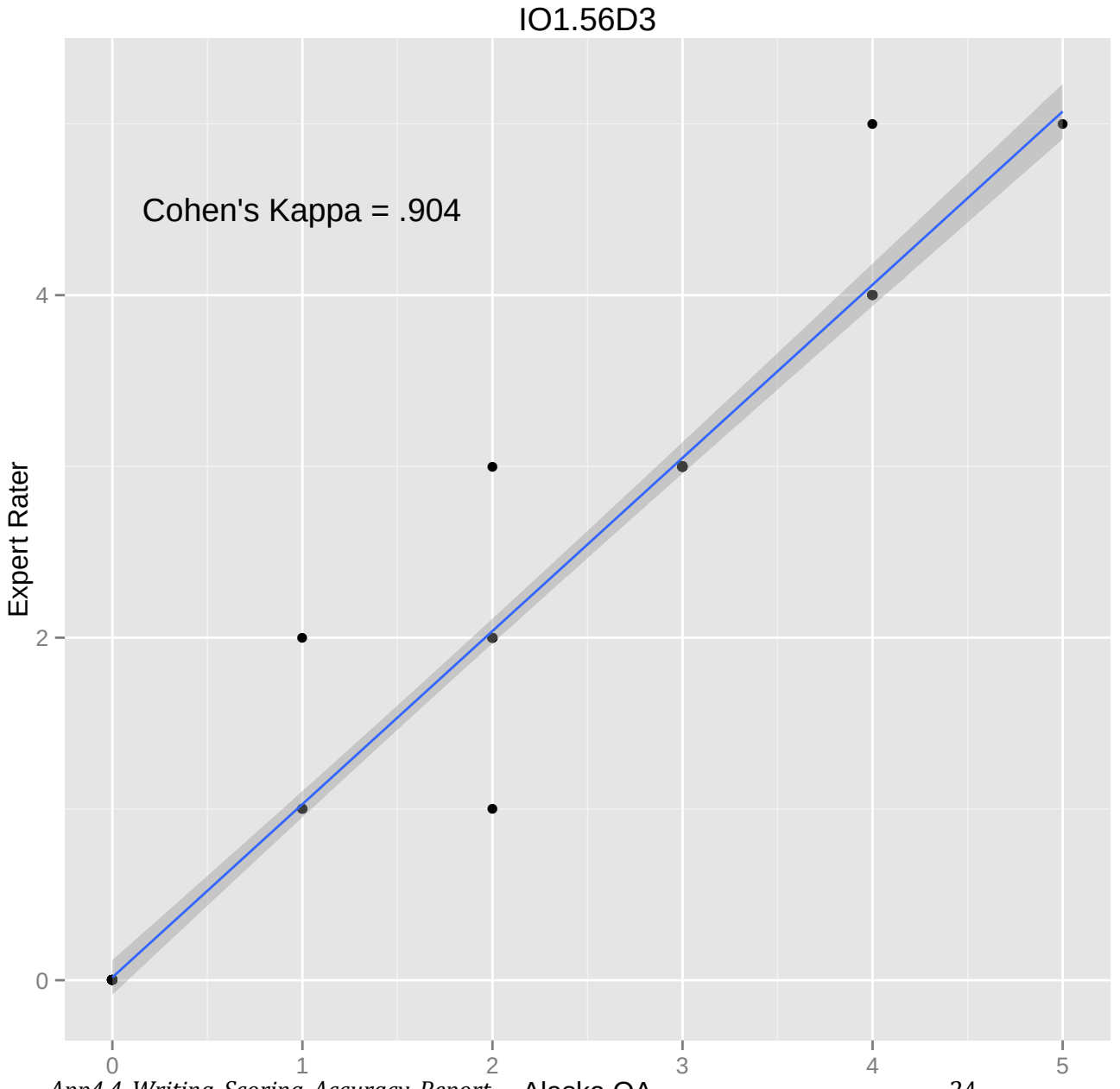

IO1.56D2
Cohen's Kappa = .881
Expert Rater
Alaska QA
0
2
4
0
1
2
3
4
5


1.56D3num
Cohen's Kappa = .970
Expert Rater
7.5
5.0
2.5
0.0
0
5
10
15
20
25
Alaska QA


1.56D3den
Cohen's Kappa = .953
Expert Rater
Alaska QA
25
20
15
10
5
0
0
10
20
30
40


CWS1.56D3pct
Cohen's Kappa = .913
Expert Rater
Alaska QA
100
75
50
25
0
0
25
50
75
100


CWS1.56D3pts
Cohen's Kappa = .940
Expert Rater
Alaska QA
10.0
7.5
5.0
2.5
0.0
0.0
2.5
5.0
7.5
10.0


IO1.56D3
Cohen's Kappa = .904
Expert Rater
0
2
4
0
1
2
3
4
5
Alaska QA